

Probability analysis in diagnosing coronary artery disease

Citation for published version (APA):

de Korte, P. J. (1993). *Probability analysis in diagnosing coronary artery disease*. [Doctoral Thesis, Maastricht University]. Datawyse / Universitaire Pers Maastricht. <https://doi.org/10.26481/dis.19930311pk>

Document status and date:

Published: 01/01/1993

DOI:

[10.26481/dis.19930311pk](https://doi.org/10.26481/dis.19930311pk)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

PROBABILITY ANALYSIS IN DIAGNOSING CORONARY ARTERY DISEASE

ISBN 90 5278 060 9

Lay out: Thum Aarts, Epidemiologie RL, Maastricht
Productie: Datawyse | Universitaire Pers Maastricht

Probability analysis in diagnosing Coronary Artery Disease

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Rijksuniversiteit Limburg te Maastricht,
op gezag van de Rector Magnificus, Prof. Mr M.J. Cohen,
volgens het besluit van het College van Dekanen,
in het openbaar te verdedigen
op donderdag, 11 maart 1993 om 16.00 uur

door

Philippus Jacobus de Korte

geboren te 's Gravenhage in 1939

Promotores:

Prof.dr F. Sturmans
Prof.dr J.M.A. van Engelshoven

Beoordelingscommissie:

Prof.dr P.J. Brombacher (voorzitter)
Dr F.W.H.M. Bär
Prof.dr J.M.J.P. van der Linden
Prof.dr O.C.K.M. Penn
Prof.dr J.H.J. Ruys (Katholieke Universiteit Nijmegen)

aan mijn ouders
Amelia
Jurjen

Contents

page

1	Introduction	9
1.1	Stemming the rise in national health care expenditures	9
1.2	Optimising the diagnostic work-up for suspected coronary artery disease	10
1.3	Aim of the study	11
2	Objectives, design and data collection of present study	13
2.1	Objectives and design	13
2.2	Variables included in the probability analysis	14
2.3	Outline of the study	15
2.4	Study population and data collection	16
2.4.1	The sampling	16
2.4.2	Type of data collected	17
2.5	Reservations about study design and collection of data	20
3	Cine coronary arteriography: the gold standard, for the diagnosis of coronary artery disease	23
3.1	Indications for cine coronary arteriography	23
3.2	Risks	24
3.3	Radiation risk	24
3.4	The definition of a lesion of haemodynamic importance	25
3.5	Grading of lesions	26
3.6	Scoring and reporting of the coronary arteriogram	27
4	Principles of probability analysis for developing diagnostic protocols	31
4.1	Parameters for describing the performance of a diagnostic test, sign or symptom	31
4.2	Sensitivity and specificity	31
4.3	Receiver operating characteristic curve, a ROC curve	33
4.4	Likelihood ratios	34
4.5	Sources of bias	36
4.6	Dependence between diagnostic variables	37
4.7	Odds Ratios	38
4.8	Collinearity and interaction	40
4.9	Logistic regression as method for estimating an OR	44
4.9.1	Introduction	44
4.9.2	The logistic function	45
4.9.3	Logistic function and Odds Ratio	46
5	Determination of $P_{\text{threshold}}$ indicating when an arteriogram should be performed	51
5.1	$P_{\text{threshold}}$	51
5.2	Determination of $P_{\text{threshold}}$ by means of a medical decision analysis	51
5.3	Determination of $P_{\text{threshold}}$ based on the budget	54
5.4	Conclusion	55
	Appendix: log-linear regression model	63

	page
6 Arteriographic findings, their distribution over sex, age and symptoms and the diagnostic value of these clinical variables in the diagnosis of coronary artery disease	67
6.1 Literature review of the value of the clinical variables sex, age and symptomatology	67
6.2 Arteriographic findings in relation to sex, age and symptoms	70
7 Electrocardiographic stress testing and its value in the detection of coronary artery disease	75
7.1 Literature review	75
7.1.1 Introduction	75
7.1.2 ST segment change	75
7.1.3 Conclusion	77
7.2 Results of stress ECG testing	77
7.2.1 Likelihood Ratio according to degree of ST segment depression	77
7.2.2 Influence of clinical variables	80
7.2.3 Probability analysis	82
7.2.4 Possible Bias by selective referral	83
7.2.5 User of Beta-blockers	83
7.2.6 Conclusion	84
8 The value of myocardial perfusion scintigraphy in the diagnosis of coronary artery disease	87
8.1 Literature review	87
8.2 Results of thallium scintigraphy	89
9 Coronary calcifications and coronary artery disease	93
9.1 Literature review	93
9.1.1 Conclusion	96
9.2 Results of (cine) fluoroscopy	97
9.2.1 Likelihood Ratio for coronary artery calcification	97
9.2.2 Influence of clinical variables	98
9.2.3 Probability analysis	101
9.2.4 Conclusion	102
10 Literature review of proposed protocols: the need for additional information	103
10.1 Literature review	103
10.2 Limitations of proposed protocols and need for additional information	107

11	Construction of diagnostic flow charts	111
11.1	Assuming independence between diagnostic variables	112
11.2	Accounting for dependence between diagnostic variables	113
11.3	Referral for non-invasive tests prior to arteriography	115
11.4	Criterion for non-invasive tests	115
11.5	Determination of the optimal sequence of tests	117
11.6	Determination of the cost-effectiveness ratio of a non-invasive diagnostic step	119
	Appendix 1: Estimation of the costs of the different non-invasive tests and cine coronary arteriography	124
	Appendix 2: The averaged benefit of performing one non-invasive test expressed in B_{therapy} units	125
	Appendix 3: Flow charts	128
	Appendix 4: Summary of proposed protocol	132
12	Evaluating a protocol in practice	133
12.1	Introduction	133
12.2	Composition of the study population for evaluating the protocol: arteriographic findings, their distribution over gender, age and symptoms	136
12.3	Results of the application of the protocol	138
12.4	Some remarks with respect to the interpretation of the results of the protocol	142
12.5	Conclusion	143
13	Validation of simple fluoroscopy with cine fluoroscopy as detection mode of calcification in the coronary arteries	145
14	Summary and conclusions	149
14.1	Aim of the thesis	149
14.2	Conclusions in regard to the method used for the probability analysis	149
14.3	General conclusions regarding the protocol in the diagnosis of coronary artery disease	150
14.4	Some specific aspects of the protocol	151
14.5	Conclusion	152
15	Samenvatting	153
16	List of references cited	157
	Acknowledgement	167
	Curriculum Vitae	171

1 | Introduction

1.1 Stemming the rise in national health care expenditures

There has been mounting pressure on the medical profession in recent years to participate in stemming the rise in national health care expenditures. As limitations on health care resources increase, it becomes essential to try to reduce the number of diagnostic and therapeutic actions by reviewing the indication criteria and thus improving the cost effectiveness of these interventions.

Already in 1978, Enthoven (1978)^{E1} indicated that measures would be necessary "to cut the costs without cutting the quality of care". He stated that "physicians should become involved in cost-effectiveness analysis for medical decision making. A synthesis of principles of economics, statistics, probability analysis and decision theory should be applied to the complex and uncertain problems of decision making. The goal is to elucidate all the costs, risks and benefits of alternative courses of action so that decision makers can be well informed in applying the necessary judgments. Such an analysis should be an aid to judgment, not a substitute for it". The aforementioned principles have already been brought to our attention by McNeil (1975)^{M1}.

The medical profession will be increasingly called upon to account for its utilisation practices (Maynard, 1987^{M2}; Appleby, 1987^{A2}; Heiberg, 1987^{H1}). These vary from place to place in a surprising fashion and as Larson (1985)^{L1} and Ruijs (1985)^{R1} stated, more research is indicated.

As Dutch hospitals are feeling the pressure of public spending cuts the hospital management is now engaged in elaborating a budgeting method which has been obligatory since 1983 for all hospitals. Under the old system, all hospital costs were covered by the patient care or "nursing" fees, a typical open-end system.

What happened, more or less, was that hospitals simply aimed at maximising the services provided. This was one of the main reasons why the costs of Dutch health care have skyrocketed to Dfl. 34.4 billion in 1985, Dfl. 35.8 billion in 1986 and Dfl. 36.7 billion in 1987, which is about 8.4% of the country's gross national product. The costs for 1988 and 1989 were Dfl. 37.5 and Dfl. 39.3 billion, respectively.

The new funding system obliges hospitals to work according to a budget. All care has to be provided within this fixed budget. Physicians' fees, as yet, are not normally included in the budget, but all costs generated by the physicians, e.g. diagnostic investigations, are included. As a consequence of this new external

budgeting system, hospitals are now introducing an internal budgeting system for the different departments.

This also applies to the diagnostic radiology department where it is of importance to optimise the diagnostic work-up. Apart from the economic aspect, unnecessary investigations can be inconvenient and hazardous for the patient, especially invasive tests such as arteriography. Furthermore, unnecessary investigations may provide contradictory results, impeding effective decision making.

In the Netherlands, a Committee, installed by the Dutch Government (Committee Choices in Health Care^{D1}), very recently concluded that it is necessary to take decisions concerning priorities in health care and various criteria have been proposed in order to analyse the value of a health care provision. Firstly, the health care provision should be necessary; secondly, it should be effective or, stated differently, it should be demonstrated that the provision does have an influence on the condition of the patient. Thirdly, it should be efficient, which can be determined by a cost-effectiveness analysis or a cost-utility analysis, as proposed by the Committee. Finally it is proposed to make an appeal to the responsibility of the citizens themselves: provisions which can reasonably be paid for by the citizens themselves should not be insured. The kind of analysis for evaluating a set of diagnostic procedures, according to the second criterion of the Committee, its effectiveness, is very complex, and in current literature, studies on this subject are not often encountered.

1.2 Optimising the diagnostic work-up for suspected coronary artery disease

A relevant subject for reviewing the indication criteria is the diagnostic work-up of patients suspected of coronary artery disease. It is generally accepted that for the definitive diagnosis of coronary artery disease cine coronary arteriography is required. However, this investigation is one of the most expensive radio-diagnostic examinations and facilities are limited. This is partly due to the necessary hospitalisation of the patient. Furthermore, the score of coronary arteriograms with significant disease can vary substantially among the clinicians. For example, in a large Dutch hospital this score varied over the period 1983-1987 from 20% to 94% (Figure 1.1).*

In the clinical evaluation of patients with chest pain, the decision to refer for coronary arteriography is a judgement based largely on the character of the symptoms and on the age and sex of the patient. The physician makes this judgement by experience and intuition, recognising that typical symptoms of angina, advancing age and male sex independently heighten the probability of coronary artery disease. Non-invasive testing (stress ECG and thallium scintigraphy) may be of further help. Furthermore, the literature indicates that the (cine) fluoroscopic detection of coronary artery calcifications is an under-used, cheap

* The percentage of normal investigations has also been used as one of the indicators in a study to evaluate the quality of the referral performance (Phibbs, 1988^{F6}). The percentage of "normal" results, i.e. no lesions of 75% cross-sectional area loss, ranged from 8 to 17% for the 4 hospitals included in his study. He concluded that the finding of a "normal" investigation in a patient presumed to have significant disease was an opportunity for postgraduate education (was the referral of this patient correct, and could this result have been expected?).

and non-invasive test for coronary artery disease and may be an alternative to the more expensive non-invasive tests in the work-up of coronary artery disease.

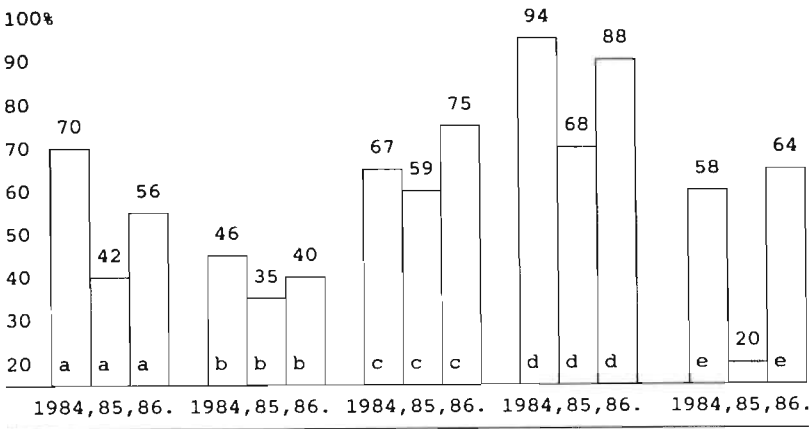


Figure 1.1 Percentage of coronary arteriograms showing significant coronary artery disease, referred by 5 cardiologists, (a,b,c,d,e), in 1984, 1985 and 1986

Timmis (1985)^{T1} discussed the possibility of optimising this diagnostic process in the treatment of coronary artery disease, especially the selection process for coronary arteriography, by using a protocol. Two problems arise in the development of such a protocol. Firstly, a protocol will be based on a probability analysis which deals with the probability of having disease given a certain clinical history, eventually combined with non-invasive test results. The current way of analyzing a diagnostic procedure is by sequentially using the theorem of Bayes for each diagnostic step, assuming that each diagnostic variable estimates the disease probability independently of the results of the foregoing diagnostic steps (Sackett, 1985^{S1}). As judged by the literature the manner in which a probability analysis should be performed is a point for discussion.

Secondly, with help of the results of a probability analysis it is possible to weight and compare the clinical outcomes of the different alternative diagnostic procedures and select the optimal procedure. The standard ways to assess these outcomes are cumbersome and would take many years. For our purposes it will be necessary to use a more simple and straightforward method.

1.3 Aim of the study

The primary goal of this study is the development of a protocol for the diagnostic work-up for patients suspected of coronary artery disease. This protocol should indicate which patients referred for arteriography should undergo this investigation and which, if any, non-invasive investigation(s) have to be performed in advance to substantiate this decision. The expected advantages of such a protocol over the existing diagnostic work-up would be to give priority to those patients who need a cine coronary arteriogram more urgently than others, as well as cutting costs by avoiding unnecessary non-invasive testing. Special attention will be paid to the evaluation of (cine) fluoroscopy as a non-invasive test.

The methods to approach the problems with the probability analysis and assessment of the clinical outcomes as used for this protocol could serve for the development of a more general method in analyzing the effectiveness of a diagnostic health care provision, as proposed by the Dunning Commission.^{D1} Whether the resulting protocol has to be considered as an exercise of this approach for protocol development or can directly be applied in medical practice needs to be determined by the results of a prospective study. In the last part of the present study the outcome of normal (unaffected) clinical practice has been compared with the consequences of applications of the proposed diagnostic protocol for patients suspected of coronary artery disease.

Notation of references

The bibliographic references are denoted by the first letter of the last name of the first author and a number, such as L¹, rather than simply by a number. This convention combines the convenience, simplicity, and brevity of a reference symbol in the text with the usefulness to the reader of being able easily to find the reference to a given author's paper in the chapter of references, which are primarily arranged alphabetically. The secondary, numerical arrangement has no logic, and new references within each alphabetical subheading are simply added to the end of the list in that subheading without the need for renumbering all subsequent references.

This method, taken from Kirklin & Barratt-Boyes, greatly simplified the preparation of this thesis and is recommended to others (Kirklin & Barratt-Boyes, 1986).^{K1}

2

Objectives, design and data collection of present study

2.1 Objectives and design

The objective of the present study was the construction of a protocol to optimise the diagnostic process for patients suspected of having coronary artery disease. Such a protocol should answer four questions:

- a. When should a cine coronary arteriogram be performed? In other words, for which patients is arteriography indicated?
- b. Which, if any, non-invasive test should be performed for an optimal referral for arteriography?
- c. If more than one test needs to be performed, what is the optimal sequence?
- d. Is it possible to quantify the ratio between costs and effectiveness of a diagnostic step in such a way that this can be used for the decision whether or not to perform the test?

a. When should a patient be referred for arteriography?

Whether patients with certain diagnostic characteristics need referral for arteriography is ultimately determined by the average clinical improvement of the patients as the result of such a referral. Hypothetically, this could be assessed by performing an experiment. In such an experiment a number of patients would be divided at random into two equal groups. Patients in one group are referred for arteriography whereas this is not performed in the other group. The results of these two groups are compared after a certain period, as judged by survival, QALYs (quality adjusted life years) etc. Differences determine the value of arteriography. In a clinical setting, however, this approach is not possible. It will be unacceptable to the cardiologist and it will not be feasible as the many results needed require a large study population and a long evaluation period.

An alternative method would be to use a medical decision analysis or cost-effectiveness analysis. Although this would have been possible yet another method has been chosen.

For reasons already discussed in Chapter 1, currently, the board of directors of the hospital sets the number of coronary arteriograms which may be performed each year. This depends on the budget of the hospital, which is determined by negotiation between the management of the hospital and the insurance companies, and which can be viewed as a consensus of most of the parties involved. Now, how can the number of coronary arteriograms permitted help to decide which patients should be referred?

Several criteria, such as symptomatology or age, could be used to select the patients with the highest indication for coronary arteriography.

In the following the referral pattern leading to the highest score of arteriograms with significant coronary artery disease will be considered as the optimal one. Thus, the probability of having coronary artery disease will determine if an arteriogram is performed and therefore for all patients this probability of finding significant coronary artery disease has to be estimated. This estimation will be based on those diagnostic variables which are judged relevant in the literature. Referral of patients with the highest probability of having coronary artery disease will improve the score. But, as the number of coronary arteriograms which may be performed has been frozen, there has to be a cut-off point or threshold that determines this referral decision. This means that the patient will have an arteriogram if the probability is higher than this threshold, called $P_{\text{threshold}}$, and the patient will be denied an arteriogram if the probability is less than $P_{\text{threshold}}$. The directives of the hospital management as a consequence of the compulsory budget were used to establish the $P_{\text{threshold}}$. This will be further elucidated in Chapter 5.

It may be that the $P_{\text{threshold}}$ determined in this way appears to be too high while the cut-off point is in part arbitrary. In this case one has an instrument to substantiate the claim that the number of arteriograms allowed is too low, implying that the budget ought to be increased to provide the necessary medical care.

b. Which non-invasive test should be performed for the optimal referral for arteriography?

The protocol should advise, based on history taking, which non-invasive test, if any, should be performed in order to select each year those patients for arteriography who have the highest chance of coronary artery disease. The decision analysis could be used here also, but the $P_{\text{threshold}}$, as discussed in the previous paragraph, also will be used.

If the pre- and post-testing probabilities for a certain patient category are both higher or lower than the $P_{\text{threshold}}$, the non-invasive testing does not influence the referral decision. Thus, the $P_{\text{threshold}}$ will decide if non-invasive testing is recommended.

c. If more than one test is to be performed, what is the optimal sequence?

For those patients in whom more than one non-invasive test changes the referral advice, the optimal sequence has to be determined. It is reasonable to expect that the less costly and better discriminating test will be used first.

d. What is the cost-effectiveness ratio of a non-invasive test?

The last question that has to be answered is whether the ratio between costs and effectiveness of a diagnostic step can be quantified and can be used for the decision whether or not to perform the test.

These last three questions are further elaborated in Chapter 11.

2.2 Variables included in the probability analysis

There are many diagnostic variables which may be included in the probability analysis. In this study only those variables are included which have been judged in the literature as the important ones: three based on history (sex, age and

symptoms) and three non-invasive tests (stress ECG, thallium scintigram, and (cine) fluoroscopy). There are other data which could be used by the cardiologist in his decision to refer the patient but are of less importance as it seems from the literature. These diagnostic variables such as the ECG at rest, smoking habits, blood cholesterol levels, weight, family history, etc., have not been included in this analysis.

In this analysis special attention will be given to (cine) fluoroscopy, as the method for the detection of calcification in the coronary arteries. As judged by the literature, this is an under-used, non-invasive test for coronary artery disease. Cited advantages of its use include a good diagnostic yield, patient convenience, independence from medication and exercise capacity (in contrast to stress ECG testing) and last, but not least, its low cost.

2.3 Outline of the study

The presentation starts by discussing the cine coronary arteriogram, the gold standard, the reference test, for the various non-invasive tests utilised to diagnose coronary artery disease. Coronary arteriography is still the only method by which coronary artery disease can be diagnosed with certainty. Various parameters assessing the diagnostic value of a test are reviewed. The diagnostic odds ratio (OR) is introduced as the test parameter used in this study. Reasons for using logistic regression for a correct probability analysis are discussed.

The literature has been scanned for protocols aiding the diagnostic process of coronary artery disease. It appears that the degree of probability of having coronary artery disease determines which patient will have an arteriogram. The estimated probability is based on various diagnostic variables. The probability, or threshold, which decides whether or not an arteriogram will be performed appears to be chosen arbitrarily. The diagnostic variables used to estimate the probability of disease differ from protocol to protocol, especially with regard to the non-invasive tests used.

The literature review part is finalised, in concluding that questions may be raised regarding the methods used for the probability analysis of the published protocols. The estimates of post-test probabilities assume independence to be present between the various diagnostic tests. The implications of the presence of dependence on post-test probability estimates is not evaluated in these studies. Costs of the various non-invasive tests are also not considered in the final choice of protocols.

With help of the literature, a review has been made of the discriminating value of a number of diagnostic variables, such as age, sex and symptoms, the stress ECG, the thallium scintigram, and (cine) fluoroscopy.

In a probability analysis accounting for dependence between the diagnostic variables, one needs information about sex, age and symptoms as well as the results of the different tests in each individual patient. The data provided in studies published to date are insufficient and are of no use for our aim.

We have therefore chosen to formulate our own protocol based on findings derived from our own study population, patients seen at the De Wever Hospital in Heerlen.

Whether or not dependence exists between clinical variables and a non-invasive test and whether or not this dependence influences the post-test probability

estimate, are evaluated for stress ECG testing (Chapter 7), thallium scintigraphy (Chapter 8) and (cine)fluoroscopy (Chapter 9).

Determination of $P_{\text{threshold}}$ indicating when an arteriogram should be performed is addressed in Chapter 5. The different steps in the construction of a diagnostic flow chart are described in Chapter 11, resulting in the definitive flow charts.

As the last step, the resulting protocol accounting for dependence has been prospectively evaluated.

A second important reason for a prospective evaluation study is the fact that the cine coronary arteriogram report, thus cine fluoroscopy, was the basis of our data collection for the detection of calcifications. This method differs from "simple" fluoroscopy, the usual way of performing this non-invasive examination in clinical practice. As the detection of calcifications seems to play an important role in the protocols as constructed for patient selection, it is essential to know if the discriminating value of simple fluoroscopy differs from that of cine fluoroscopy. A prospective study is the only possible approach to answer this question. Thus, the protocol is evaluated prospectively, making it necessary to use a new study population. The thesis ends with some comments regarding the advantages and disadvantages of using a protocol, together with a summary and some final recommendations.

2.4 Study population and data collection

2.4.1 The sampling

During the period from January, 1983 to March, 1987, a total of 1921 cine coronary arteriograms were performed and reported. All records of the patients involved could be traced. 833 patients of this group were referred for arteriography to rule out or confirm coronary artery disease by six cardiologists practising in an 800-bed teaching hospital. Excluded were all other referred patients for whom the diagnosis of coronary artery disease was already proven (e.g. a previous myocardial infarction or an abnormal coronary arteriogram) or for whom the study was incidental in the evaluation of congenital or valvular heart disease and where coronary artery disease was not the main issue. After reviewing the medical records of these 833 subjects additionally 55 patients were also excluded in whom the diagnosis of unstable angina was made by the criteria described in the next section. These patients were excluded from this study because the referral for arteriography is not questionable while non-invasive testing is usually not indicated and thus not performed or may even be medically contraindicated; the remaining 778 patients formed the study population. Figure 2.1 demonstrates schematically the selection of the study population.

The sampling was partly retrospective and partly prospective and started in May, 1986. No data were collected from patients before January, 1983, as the number of coronary arteriograms differed too much with subsequent years. This was likely due to a different patient work-up and a different referral pattern for arteriography. At the ending of the sampling period the study population included at that moment 778 persons and the size was judged to be sufficient enough for our purpose. Furthermore, it was thought necessary to begin the analysis of the data in order to be able to supply possibly useful information for the current discussion of the problem as stated in the introduction.

2.4.2 Type of data collected

In 778 patients age, sex, symptomatology and the results of the stress ECG, thallium scintigraphy and cine fluoroscopy for arterial coronary calcification and the cine coronary arteriography (details given in the next sections) were recorded. The number and sequence of the non-invasive tests (stress ECG and thallium perfusion scintigraphy) varied from patient to patient. No attempt was made to influence the referring cardiologist in his choice of diagnostic tests or in his decision to recommend coronary arteriography. Consequently not all tests were performed on all patients and the resultant data are, as already noted, incomplete. Whether or not the pattern of non-invasive testing and of referral for coronary arteriography was also representative of other hospitals was not addressed by our analysis. Data older than 12 months as compared with the date of the cine arteriogram were omitted.

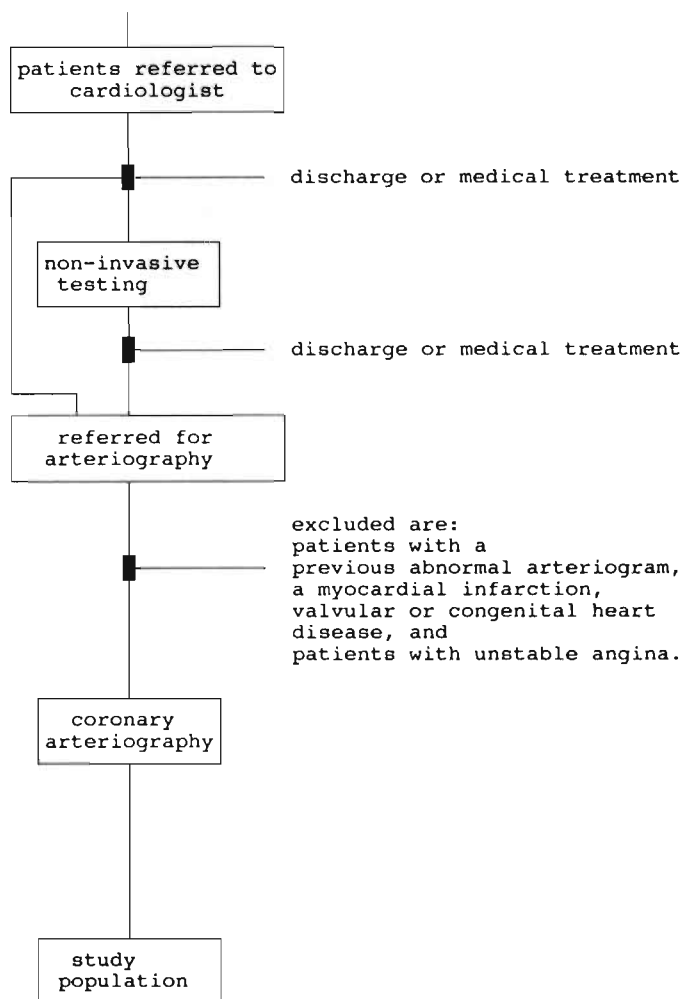


Figure 2.1 Selection process for the study population which provided the data for the protocol

Symptomatology

The symptoms of the patients were classified on the basis of the documented history taken by the referring cardiologist.

Definitions for symptomatology.

- (1) Typical (stable) angina: retrosternal chest discomfort or pain precipitated by exertion and other recognised causes, promptly relieved by rest or nitroglycerin (within 5 minutes).
- (2) Atypical angina: retrosternal chest discomfort or pain with some but not all of the above features (discomfort which is either not retrosternal, not precipitated by exertion or other recognised causes or not relieved by rest or nitroglycerin).
- (3) Nonspecific chest pain: chest discomfort or pain with none of the above factors present.
- (4) Asymptomatic: no chest pain or discomfort (this category includes patients with arrhythmias, abnormal ECG findings at rest, signs of left ventricular failure etc.).
- (5) Unstable angina: retrosternal chest discomfort or pain occurring with or without previous typical angina, occurring at rest and/or of recent onset. Unstable angina was diagnosed if the symptomatology previously categorised as typical angina changed in character over a short period, and if the pain was more easily provoked or present at rest and/or was no longer promptly relieved by rest or nitroglycerin. (Unstable angina is a complex condition, not easy to define, as discussed in a recent article by Braunwald, and probably consists of a number of conditions, all characterised by severe transient myocardial ischaemia; several other terms have been used for this condition (Braunwald, 1989)^{B1}.)

Stress ECG

Stress ECG testing as performed in our hospital utilises bicycle ergometry. Exercise is begun with a work load of 50 Watts and the load is increased by 10- to 20-Watt increments at 1.0 minute intervals. Test end-points are (1) progressive angina (2) ST depression greater than 2 mm (3) inability of the patients to continue due to physical exhaustion (4) a heart rate reaching 85% of the maximum predicted for the subject's age and sex. An electrocardiogram is recorded using 12 leads before the start of the test as well as at its finish. Three leads (V1, V5 and AVF) are continuously monitored and recorded on paper every minute during the exercise and during the first 5 minutes of recovery. Cuff blood pressure measurements are obtained every two minutes. The ECG data are evaluated and reported by the referring cardiologists.

For our study the findings of the stress ECG were extracted and recorded from the patient data, while the result of the coronary arteriogram was unknown. Only the response of the ST segment was used to classify the test result.

Any junctional depression in addition to the ECG at rest followed by a horizontal or downsloping ST segment of at least 80 msec duration after the J point was read as a positive response. A junctional depression followed by a slowly upsloping ST segment of at least 2 mm, 80 msec after the J point was also read as positive. The magnitude of ST segment depression at J point was recorded in 0.5 mm intervals ranging from 0.5 mm to over 2.5 mm depression.

A negative response in the stress ECG was recorded if no ST segment abnormality occurred with adequate exercise. A heart rate of 85% maximum, with correction for age and sex, was judged to be adequate. The test was also read as negative if the ECG at rest was abnormal and no ST segment depression occurred in spite of adequate exercise.

The stress ECG was recorded as non-diagnostic in the absence of any ST segment depression but an inadequate exercise level (with heart rate of less than 85% of the maximum predicted for the subject's age and sex).

The stress ECG was reviewed "blindly" by another (always the same) cardiologist, if there was any doubt about the result. His opinion prevailed.

(Assessment of the presence or characterisation of pain during the test was not recorded for this study, as the tests were performed by personnel of variable experience which could have led to unreliable observations.)

Thallium scintigraphy

The examination as performed in our hospital is as follows.

The patient undergoes exercise by bicycle ergometry, according to a standardised protocol with administration of successively increasing workloads until either symptoms, angina or fatigue, limit further exercise or the patient has achieved at least 100% of age-predicted maximum heart rate. 2 mCi of thallium-201 is intravenously injected at that point and the exercise is continued for another 2 minutes at reduced exercise load. Thallium imaging is then started, using a low-energy, all purpose (LEAP) collimator and a 0.5-inch crystal. Images in left anterior oblique, anterior and left lateral projections are obtained with approximately 300,000 counts and at least 1000 counts / cm² at the point of highest concentration. Redistribution images are obtained in the same projections four hours later. The images are viewed using a computer video display, allowing contrast and brightness adjustment. The computer also permits a rapid determination of the heart-lung ratio as well as background subtraction. No further quantitative data such as wash-out curves are obtained at this data sampling period. All the images are recorded on film. The test is recorded as normal if there are no defects reported (by the physician of the department of nuclear medicine performing the study) and if the exercise is adequate (reaching a heart rate of at least 85% or more of the predicted maximum heart rate). It is recorded as "fixed abnormality" if defects observed during exercise persist during the redistribution phase and is recorded as "reversible abnormality" if defects present during exercise show significant correction during the redistribution phase. Both these findings are recorded as a positive test, without further differentiation.

The patients with a fixed defect, suggestive of infarcted tissue, were not eliminated from the study population, in view of the lack of certainty of this finding with the present technique (Pohost, 1990).^{P1}

The test was classified as non-diagnostic if no defects were observed with inadequate exercise.

Arterial coronary calcification

As simple fluoroscopy is not routinely performed in the Netherlands, the cine coronary arteriogram served both to diagnose calcification in the coronary arteries and to confirm the presence or absence of coronary artery disease. The presence

and localisation of any calcification present and whether lesions of haemodynamic significance involved calcifications, were also recorded.

All reports on cine coronary arteriograms mention the presence or absence of calcification in the coronary arteries, as recorded on the cine film during the coronary cine arteriogram.

Cine coronary arteriography

Cardiac catheterisation was performed by standard methods, using the Judkins technique and dedicated angiographic equipment. An LV angiogram was obtained in 40-degree, right anterior oblique projection, in some cases with an additional 30-degree, left anterior oblique projection. A tablet of Nitrostat was given routinely before opacification of the coronary arteries. Selective coronary arteriography was performed with multiple projections of the left (an AP and left lateral projection, a number of LAO and RAO projections routinely made, as well as projections with caudo-cranial tilt) and right coronary artery (at least two, usually three, projections). The opacification of the coronary arteries was recorded on cine film using a standard 6.5-inch (15-cm) image intensifier.

Philips radiographic equipment was used, consisting of an Optimus M200 generator, a cesium iodide input phosphor 15-/ 22.5-cm image intensifier. The 35-mm film resolution of the 15-cm intensifier, used during filming of the coronary arteries, was 25 lp/cm.

All the cine coronary arteriograms were examined by one of a team of 2 radiologists, both specialised in cardio-radiology. The cine coronary arteriograms were reported and scored according to the system published by Brandt and will be discussed in Chapter 3 (Brandt, 1977).^{B2} This system utilises not only an estimate of the severity of stenosis in a vessel, but also an estimate of the importance of the vessel according to the proportion of myocardial tissue supplied. It further takes into account the presence of more than one lesion in the same vessel, as well as the presence of lengthy lesions.

Significant coronary artery disease was defined in our study as the presence of at least one coronary artery with a lesion of 75% loss in cross-sectional area, supplying at least 1.5 myocardial units. The value of the myocardial score was extracted from the clinical notes on each patient.

Furthermore, the patients were divided into 5 categories (1 - 5), with regard to the coronary arteriogram findings:

- (1) no coronary artery disease at all,
- (2) no coronary artery disease, but only a muscle bridge,
- (3) only minimal coronary artery disease, no significant stenosis present,
- (4) coronary artery disease, with at least one lesion of 75% loss in cross sectional area (a moderate stenosis) and the artery supplying 1.5 myocardial units, but no 90% lesions,
- (5) coronary artery disease, with at least one lesion of $\geq 90\%$ loss in cross-sectional area (a severe stenosis), and the artery supplying 1.5 myocardial units.

2.5 Reservations about study design and collection of data

The ideal study design would be to perform the cine coronary arteriogram, the gold standard, and the diagnostic tests to be evaluated in all patients consulting

the cardiologists with regard to symptoms or signs suggestive of ischaemic heart disease. It is, however, unethical to perform an arteriogram without a medical indication, submitting a patient to the inconvenience and the risks of such an invasive examination, and to the possible harmful effects of the radiation as well. The same holds to a lesser degree for the non-invasive tests.

This necessitated that the study population had to consist only of patients referred for coronary arteriography. This was a disadvantage, as the study population thus selected may indeed not have been representative of all patients consulting the cardiologist. If the referring cardiologist has very strict criteria for arteriography referral, only limited patient groups with special clinical characteristics will enter the study. However, patients in our study were referred by six cardiologists and it is unlikely that all had the same criteria for angiography referral. Moreover, strict adherence to clinical criteria is not likely as other factors, less related to disease than to "patient pressure", may well influence the choice in the diagnostic process. These assumptions are supported by Figure 1.1 which is based on our data. This figure shows that the percentages of coronary arteriograms showing significant disease varied between the different cardiologists as well as over the years for the individual cardiologists. It is therefore expected that our study population consists of patients with all combinations of age, sex and symptomatology (and perhaps other unknown variables). Restricting the results (and conclusions) in the first instance only to those patients for whom an arteriographic study is considered, counteracts this selection bias.

The second selection bias is due to the fact that the stress ECG and the thallium scintigram were not both performed in all the subjects taking part in the study. In contrast with the previous described bias, the impact of this bias on the results can be investigated as for patients in whom the tests have not been performed, still the results of all other diagnostic variables are known. This allows a stratified comparison of the estimate of the disease probability for patients whether or not the test had been performed. (This is further elaborated in the appendix of chapter 5).

A second problem was the data collection of coronary calcification, as fluoroscopy is not a frequently requested diagnostic procedure in the Netherlands. All information was therefore sampled from cine coronary arteriography reports. This means that the test result was not obtained independent of the result of the gold standard. The influence of the knowledge of the arteriogram report on the test result and vice versa (diagnostic review bias), forms another source of bias. However, as a consequence of this manner of data collection, the result of the cine fluoroscopy had no influence on the selection of the patients for coronary arteriography. This kind of selection bias, commonly occurring in observational studies of diagnostic procedures, did not occur.

Whether there is a difference between the results obtained using cine fluoroscopy and ordinary fluoroscopy remains to be determined. A comparison will be performed and reported.

3

Cine coronary arteriography: the gold standard for the diagnosis of coronary artery disease

Coronary arteriography is still the only method by which coronary artery disease can be diagnosed with "certainty" and the extent of the disease be evaluated (Bruschke, 1970^{B3} and Patterson, 1982^{P2}). The commonly used technique to visualise the coronary arteries is cine coronary arteriography. Cine coronary arteriography therefore serves usually as the reference test, the gold standard, for the various non-invasive tests utilised to diagnose coronary artery disease.

3.1 Indications for cine coronary arteriography

The indications for cardiac catheterisation and a coronary arteriogram vary from cardiologist to cardiologist, (Willerson, 1985)^{W1}, as well as from nation to nation (Brook, 1988)^{B4}. It varies for patients known to have coronary artery disease as well as for patients in whom the diagnosis is still not certain. Apart from patients with various categories of chest pain, evaluation by arteriography for the possible presence of coronary artery disease may also be indicated in patients with abnormal ECGs at rest or at exercise, patients with arrhythmias, or patients with unexplained left ventricular dysfunction and in whom chest pain is not present as the prevailing symptom.

It may be considered important to determine for certain by arteriography whether atypical chest pain is related to coronary disease, even if the cardiologist involved is nearly certain that no coronary artery disease is present. Patients' lives may be disrupted by the uncertainty of having disease and dealing with infrequent visits to emergency rooms and admissions to intensive care units.

There are also contra-indications to coronary arteriography for diagnosing coronary artery disease, e.g.(1) the patient is unwilling to have either percutaneous transluminal angioplasty or surgery, (2) coexisting medical problems that may shorten life expectancy, and (3) very advanced age (perhaps more than 80 years of age) (Willerson, 1985)^{W1}. The Dunning Commission,^{D1} installed by the Dutch Government in 1990 concluded that (advanced) age as such should not be an exclusion criterion for a health care provision.

Effects of age on morbidity and mortality in coronary bypass surgery remain controversial; the large clinical trials such as the Veterans Administration Cooperative Study, the European Coronary Surgery Study and the Coronary Artery Surgery Study, included patients with an average age of 50 (Satler, 1989)^{S2}. Acinapura (1988)^{A2} reported on the morbidity and mortality of bypass surgery in 685 patients older than 70 years. Mortality was higher, 7% versus the expected mortality rate of

1%-3%. Naunheim (1987)^{N1} reported on the results of bypass surgery in 23 patients aged 80 years and older. The mortality rate was 21% and the risk of significant complications 50%.

3.2 Risks

Cine coronary arteriography is an invasive examination with an inherent, albeit rather small, risk when performed by experienced operators aware of the hazards and capable of preventing complications.

Adams (1979)^{A3} has reported a mortality rate of 0.14%, a myocardial infarction rate of 0.18% and a thrombo-embolic complication rate of 0.75%. Davis (1979)^{D2} reported based on the CASS study, using Judkins technique, a mortality rate of 0.14%, a myocardial infarction rate of 0.22%, a stroke rate of 0.08%, an arterial thrombosis rate of 0.24%, and an arterial dissection rate of 0.13%.

Patients who have clinical signs of congestive heart failure at the time of examination have a higher rate of complications; the risk of death increases fivefold.

Patients with left main disease (stenosis with loss of 75% or more cross sectional area) have also a higher mortality rate (0.75%), a rate nearly 7 times higher than that for patients without left main disease. The occurrence of angina pectoris in the 24 hours before catheterisation is another factor increasing the risk of complications (up to 31%) during the catheterisation or in the 24 hours thereafter (Gordon, 1987)^{G1}.

3.3 Radiation risk

Cine coronary arteriography involves radiation and exposure to radiation of the patient should also be of concern.

Ionising radiation causes deterministic and stochastic effects in irradiated tissue. Deterministic effects are the result of cell killing resulting in a loss of organ function. Stochastic effects occur when an irradiated cell is modified but still not killed. Such cells may eventually develop into malignant tumour cells. A hereditary effect can also occur when gene mutation and chromosomal aberrations are induced in parental germ cells resulting in genetic disorders. The International Commission on Radiological Protection (ICRP) plays an important role in the discussions about radiation protection. Principal recommendations and new concepts were published in the 1990 Recommendations (ICRP, 1990)^{I1}. On the basis of recent material on studies of survivors of the atomic bombs on Hiroshima and Nagasaki in 1945, the estimated risk of induction of malignant tumours due to irradiation is now approximately 4 times as high as had been estimated 15 years ago.

In view of this, the ICRP has recommended lowering the dose limits in occupational exposure (radiological workers). In 1990 the ICRP recommended an occupational dose limit, related to stochastic effects, of 20 mSv per year, averaged over defined periods of 5 years with the further provision that the effective dose should not exceed 50 mSv in a single year. In 1977 the occupational dose limit for stochastic effects was 50 mSv per year. The overall applied reduction factor of 2.5 reflects concern.

In regard to control of public exposure the effective dose limit in a year was reduced from 5 mSv to 1 mSv in 1983.

An indication that authorities want to reduce radiation doses comes from the National Radiological Protection Board of Great Britain which recently (1988)^{N2} issued a new recommendation that individual radiation workers should not be exposed to more than an average of 15 mSv per year, a reduction to less than a third of the current limit of 50 mSv. As Orr stated there appear to be greater risks from radiation than predicted before (Orr, 1988)^{O1}.

The Dutch authorities also aim to reduce the currently permitted levels for occupational and public exposure according to a recent proposal of the Council of Ministers dated March, 1990 (NRC, March 1991)^{N3}.

The ICRP does not recommend dose limits to patients to control the medical exposure. If a diagnostic examination is justified and protection is optimised, the dose is considered to be as low as is compatible with the medical purposes.

However, out of interest, it may be useful to compare the patient exposure during cine coronary arteriography with some other radiological diagnostic studies. Kicken compared the exposure-area product XAP, as an easy to measure dosimetric quantity which is explained in a Publication of The National Radiological Protection Board dated 1986. The mean XAP for cine coronary arteriography is 5000 R.square cm, per examination. The mean relative contribution of fluoroscopy is 40% per examination. Comparison of the mean XAP for cardiac procedures to the mean XAP for various diagnostic procedures indicates that one cardiac procedure equals approximately 120 chest X rays, 10 abdomen, 5 lumbar spine and 1.5 barium enema examinations (Kicken, 1988)^{K2}. Thus, the radiation exposure of patients obtained during a cine coronary arteriogram is relatively high.

Beentjes used a different dosimetric quantity, the averaged somatic effective dose (SED), with mSv as unit, to compare various radiological diagnostic procedures (Beentjes, 1990)^{B5}. The values for cine coronary arteriography are also the highest (Table 3.1).

Table 3.1 Averaged Somatic Effective Dose (SED), for various radiological diagnostic procedures conducted in the Netherlands (Beentjes, 1990)

Investigation	SED (mSv)
Cine coronary arteriogram	28
CT-Thorax	7
X-ray exam, oesophagus	4
X-ray exam, stomach	5
X-ray exam, colon	7
Intravenous urogram	3
X-ray exam, lumbar spine	2

3.4 The definition of a lesion of haemodynamic importance

In experimental studies, a 75% reduction in the cross sectional area of the lumen of a large artery is required to decrease flow, and such a lesion is in practice associated with a moderate stenosis, reaching haemodynamic significance (Gould, 1974)^{G2}.

Even under controlled experimental studies, there is considerable variability in response. Considering the dissimilarity between experimentally created stenosis and the multifaced presentation of coronary atherosclerosis, the extrapolations may appear inappropriate in the great majority of clinical cases (Paulin, 1987)^{P3}.

Harrison (1984)^{H2} and White (1984)^{W2} both published reports on a study correlating the arteriographic findings with intra-operative flow studies. They used a pulsed Doppler velocity probe and measured the coronary reactive hyperaemic response, which could be used to assess the vasodilator reserve of a stenosed vessel and thus provide an estimate of the functional significance of the stenosis. Their study indicated that lesions with loss of less than 75% cross-sectional area could still be significant; for example if diffuse narrowing is not recognised and is not taken into consideration when estimating the degree of narrowing. Identifying the true normal diameter of the vessel may be difficult. Optimising the identification of the true normal lumen size in patients with diffuse coronary artery disease may well enhance the ability to differentiate significant from insignificant lesions (Harrison, 1984)^{H2}. Klocke (1987)^{K3} described the use of coronary flow reserve measurements in defining the pathophysiology of a stenosed coronary artery. He concluded that at present the "complexities of performance and interpretation of flow reserve are greater than generally appreciated and that measurements are still in an early stage of clinical application and still not usable for patient care decisions."

A long lesion or serial lesions may be of greater haemodynamic significance than a short-segment stenosis of similar grade.

The terminology for coronary arterial stenoses is often ambiguous in the literature as well as in clinical reporting (Paulin, 1987)^{P3}. It is essential to state whether the percentage stenosis refers to a loss in cross-sectional area of the lumen or to a reduction in the average diameter. A 90% cross-sectional area stenosis is a lesion of 67% (or approximately 70%) average diameter reduction; a 75% cross-sectional area stenosis is one with an average 50% diameter reduction; a 50% cross-sectional area stenosis is a lesion with about 33% average diameter reduction. The term "50% stenosis" is ambiguous and fails to distinguish between, on the one hand, a 50% diameter reduction which represents a 75% cross-sectional area stenosis and is likely of haemodynamic importance, and on the other hand a 50% cross-sectional area stenosis which represents a 33% diameter reduction and is unlikely to be of immediate clinical concern.

3.5 Grading of lesions

Visual inspection of the arteriogram is currently the standard means for assessing the severity of coronary arterial stenoses (Stadius, 1990^{S3}; Johnson, 1992^{J1}). The use of an analyzing projector which allows the reporter to measure the calibre of the vessels and stenoses on the held frame can improve the objectivity and reproducibility of the assessment. However, significant intra- and inter-observer variability has been reported in the literature, consistently quoting the same three studies from the seventies (Detrie, 1975^{D3}; Zir, 1976^{Z1}; DeRouen, 1977^{D4}). However, expertise has increased and the diagnostic quality of the cine film has improved considerably since then, both factors probably reducing inter-observer variability.

In contrast, Isner (1981)^{I2} correlated angiographic and histological findings in 28 patients with left main disease and reported a remarkable observer agreement. Furthermore, he voiced the opinion that improper technique - an inadequate number of projections, the missing of at least an AP view - seemed to be the main cause of the angiographic misinterpretation of lesions.

The cross-sectional area of the lumen is not directly visible on the cine film and must be derived from observation of the diameter, preferably determined from at least two projections that profile the lesion and that are at right angles to each other, both perpendicular to the long axis of the diseased portion of the coronary artery. Because of the variable curvature and the variable complexity of the coronary arterial branching pattern, the ideal projections may not always be achieved and sometimes the assessment of severity depends on a single projection, all other views being unsuitable for profiling the lesion. A short-segment, obliquely oriented lesion may be imperfectly profiled even when the projections are perpendicular to the long axis of the appropriate arterial segment. To minimise these errors, it is essential to obtain multiple views of the coronary arteries, including appropriate cranially or caudally tilted projections when necessary. Obviously, the experience and skill of the arteriographer contribute directly to the reliability of the arteriographic assessment.

Because the grading of the severity involves comparison of the diameter at the site of a lesion with that of the normal vessel, a frequent source of potential error is the presence of diffuse narrowing (and less frequently dilatation) of the coronary arteries. A local stenosis may be interpreted as a lesion of 75% cross-sectional area loss when it should be graded 90% because it lies in a diffusely narrowed segment. This situation can usually be recognised by observing the quality of the arterial margins and taking into account the calibre of more healthy distal or neighbouring branches, as well as the size of the vascular beds supplied. The routine use of coronary vasodilators may also help to remove a possible source of error in underestimating the true normal calibre of the arteries.

Efforts to improve the accuracy of the assessment of severity of arteriographically demonstrated stenoses include digital conversion of the information in the cine arteriogram and computerised quantitative geometric and densitometric analysis (Reiber, 1981^{R2} and 1986^{R3}). Digital coronary arteriography may further improve quantitative measurements of stenoses (Vogel^{V1} in Reiber, 1986 and Zijlstra, 1987^{Z2}). These techniques however will not overcome all the difficulties discussed above, and the basic underlying requirements for an accurate assessment must include a comprehensive range of projections chosen to allow for the variable anatomy and pattern of disease of each individual coronary tree.

In conclusion, the pioneer Paulin stated in 1987^{P3} that "notwithstanding the obvious difficulties in translating anatomical lesions into their functional consequences, the coronary arteriogram, intelligently interpreted and integrated with other clinical information about a patient with suspected or manifest coronary artery disease, still remains an essential diagnostic test".

3.6 Scoring and reporting of the coronary arteriogram

Several scoring and reporting systems have been devised. The one reported by Brandt (1977)^{B2} offers certain advantages.

(1) A drawing depicting the anatomy of the individual patient's coronary arteries and the location and severity of the different lesions, collaterals, muscle bridges, presence and location of calcification, course of grafts, etc. provides easier access to relevant information than a written report alone.

As the arterial anatomy of each patient is drawn on the diagram, all anatomical variations such as a "non-dominant" right coronary artery, or a longer than usual posterior descending artery can be easily taken into account. As the anatomy of

each patient varies, systems with pre-drawn vessel anatomy, such as the system advocated by the American Heart Association (1975), are less workable. (2) The left ventricular free wall and the interventricular septum are -an arbitrary choice but based on the normal anatomy- considered to contain a total of 15 myocardial units, 7 units for the interventricular septum and 8 units for the left ventricular free wall, the latter being divided into 2 units for the diagonal territory, 3 for the obtuse marginal area and 3 for the inferior (diaphragmatic) aspect of the left ventricle. As the different vessels are drawn over these areas they are allocated an appropriate share of the myocardial units so that each vessel has a myocardial value reflecting its importance in the blood supply of the heart. Similarly, the myocardial value of the arteries supplying the interventricular septum reflects their importance and allows for the variations of each individual case. Vessels supplying 1.5 or more units are usually vessels which may be grafted if their distal quality is satisfactory. (3) The position of each lesion is indicated in the drawing as well as the severity of the lesion, estimated in percentage loss of cross-sectional area. The severity of the obstructive coronary artery disease is expressed as a myocardial score, whereby the severity of the stenosis is combined with the myocardial value of the vessel involved. A lesion of 90% cross- sectional area loss occurring in a small vessel supplying only 1 myocardial unit scores less than a 90% stenosis in a vessel supplying a larger number of units. The scoring is adapted if a lesion is longer than 1 cm or if more sequential lesions are present in the same vessel. For instance, two 50% cross-sectional area loss lesions in a row are graded as a 75% cross-sectional area loss lesion. A long 75% cross-sectional area loss lesion is graded as a 90% cross-sectional area loss lesion. Two 90% cross- sectional area loss lesions are, however, never graded as a total occlusion. Less severe lesions, as well as marginal irregularity, all contribute to the total score. The multiplying factor to calculate the myocardial score for a total occlusion is 1, for a 90% cross sectional lesion 0.8, for a 75% cross sectional lesion 0.6, etc. (See Table 3.2 for grade of severity of arterial obstruction and Table 3.3 for myocardial score.)

Table 3.2 Grade of severity of arterial obstruction, (from Brandt, 1977)

Grade	% cross-sectional area loss
a	100
b	90-99
c	75-89
d	50-74
e	< 50

Table 3.3 Myocardial score, (from Brandt, 1977)

Grade arterial obstruction	Myocardial value									
	1	2	3	4	5	6	7	8	9	10
a	1.0	2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0
b	0.8	1.6	2.4	3.2	4.0	4.8	5.6	6.4	7.2	8.0
c	0.6	1.2	1.8	2.4	3.0	3.6	4.2	4.8	5.4	6.0
d	0.4	0.8	1.2	1.6	2.0	2.4	2.8	3.2	3.6	4.0
e	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8	2.0

The sum of the individual scores gives an overall score (0 to 15) reflecting the overall severity of obstructive coronary arterial disease. (4) Particulars of the left ventricular arteriogram and other data can also be recorded on the form. The discipline in producing the individual diagram helps to maintain a high level of accuracy in analyzing the cine film and allows the written report to be much shorter than it needs to be without the diagram. The system is easy to learn and does not take more time to perform than a detailed written report. It provides an efficient rapidly comprehended communication of the information, which can be transferred to a computer for research projects.

Apart from its use at the Green Lane Hospital, Auckland, New Zealand since 1972, as reported by Kirklin and Barratt-Boyes in *Cardiac Surgery* (1986)^{K1}, it is used elsewhere in New Zealand and in some hospitals in Australia, England and Norway. It was introduced in Heerlen, the Netherlands, in 1978.

4

Principles of probability analysis for developing diagnostic protocols

4.1 Parameters for describing the performance of a diagnostic test, sign or symptom

To analyse the discriminating value of a diagnostic variable, as a diagnostic test, a sign or a symptom, one needs one or more parameters. Principles discussed in this following section apply equally to information obtained from history taking and physical examination, as well as to other kinds of diagnostic tests (Sackett, 1985⁵¹). The ideal test would be one that yields a positive or negative result with no errors. However, in medicine no such diagnostic investigations exist. Results are subject to technical and/or subjective variation and must be interpreted with knowledge of the normal range. No matter how the tests are conducted, there will be some overlap in the results between healthy and diseased persons making it impossible to separate them completely. All tests sometimes give a negative result for an individual having the condition being tested for, and sometimes give a positive result for individuals free from that condition. These results are called false negatives and false positives, respectively.

The determination of the terms for the four possible results, true positive (TP), true negative (TN), false positive (FP) and false negative (FN), is shown in Figure 4.1. It is graphically displayed in Figure 4.2, which illustrates a hypothetical distribution of a test result for two populations, one without and one with coronary artery disease. It is evident that there is a considerable change in number of patients in the four fractions TP/TN/ FP/FN, with variation of the cutoff point.

The ratio of true positives (TP) to the total number of patients with disease (TP plus FN) indicates the "sensitivity" of a test, while the ratio of the true negatives (TN) to the total number of patients without disease (TN plus FP) indicates the "specificity" (see Figure 4.1). The sensitivity and specificity of a test change if the dividing point between a negative and a positive test result is altered. For example, there is a rise in sensitivity if the criterion is lowered. Nearly all patients with coronary artery disease will now be detected but the specificity will drop. On the other hand the specificity is increased by raising the criterion level, however, with a consequent decrease in the sensitivity.

4.2 Sensitivity and specificity

The two parameters, sensitivity and specificity, are well known in the medical literature. Sensitivity or positivity in disease indicates the ability to identify abnor-

malinity; it refers to the proportion of patients with the target disorder who have a positive test result. Specificity or negativity in health defines the ability to identify normality; it refers to the proportion of subjects without the target disorder who have a negative test result.

		disease	
		present	absent
diagnostic test	positive	true positive TP	false positive FP
	negative	false negative FN	true negative TN

true positive fraction = $\frac{TP}{TP + FN}$ = "sensitivity"

true negative fraction = $\frac{TN}{TN + FP}$ = "specificity"

false positive fraction = $\frac{FP}{TN + FP}$ = 1 - true negative fraction

false negative fraction = $\frac{FN}{TP + FN}$ = 1 - true positive fraction

Figure 4.1 2 × 2 decision matrix. Determination of the terms TP, TN, FP and FN

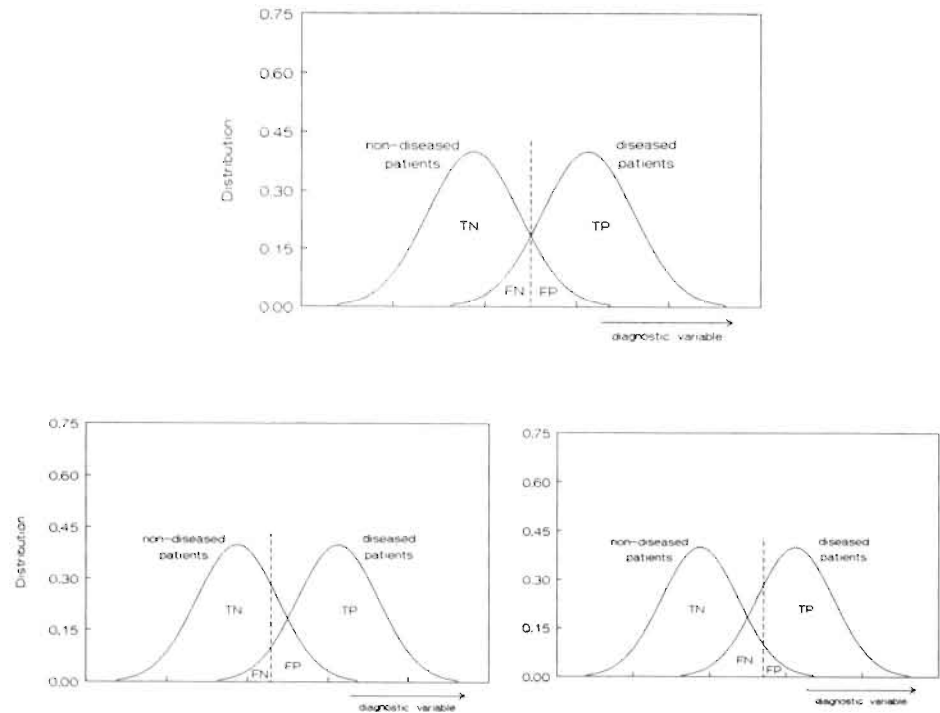


Figure 4.2 Distributions of the result of a hypothetical test for patients without and with coronary artery disease; TN=true negative; TP=true positive; FN=false negative; FP=false positive.

Only a few tests have simple binary outcomes. Most yield a continuous scale of values, of which one of several can be selected as a cutoff point for dichotomizing and differentiating into subjects with and without disease. As demonstrated, the sensitivity and specificity of a test is dependent on this cut-off point, i.e. the position of the decision line (the vertical lines in Figure 4.2) dividing the negative and positive test results. The position of this line is a matter of choice. Its position is not precisely specified for many examinations and is a matter of subjective judgment, varying from observer to observer and often from time to time with the same observer. It may be determined more easily for a test that has a numerical result, while it is more difficult to define precisely in subjective tests based on an ordinal variable. The optimal cutoff point can be chosen with the help of a receiver operating characteristic curve, a ROC curve. The different positions of the dotted lines in the figures denote the different cutoff points for this test result.

4.3 Receiver operating characteristic curve, a ROC curve

In determining the most advantageous cutoff point, it is helpful to construct a graph of the pairs of TP rate and FP rate that correspond to each possible cutoff point for the diagnostic test result. The resulting plot, a smooth, concave curve, is known as an "ROC curve" (receiver operating characteristic curve) (Figure 4.3). The location of a cutoff point along an ROC curve is called an "operating position". The operating position nearest to the left upper corner identifies the combination of the highest values of sensitivity and specificity. Only if the patient were to suffer equally from a false positive as false negative test result would this be the operating position of choice. If false positive labelling were to be harmful, it would be preferable to select a cutoff point more to the left.

ROC curves are also useful in comparing two different diagnostic tests, signs or symptoms for the same target disorder in the same test population. In the jargon of signal/detection theory, an ROC curve positioned higher and to the left has a greater inherent detectability for the signal (Turner, 1978^{T2}). However, an entirely satisfactory method has not yet been described for quantification of differences among ROC curves (Turner, 1978^{T2}).

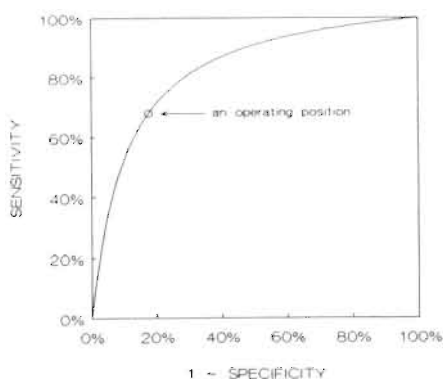


Figure 4.3 ROC curve of a hypothetical test

4.4 Likelihood Ratios

However, a disadvantage of dichotomizing is the loss of useful information. Furthermore, some test variables are not measured on an interval or ordinal scale and a ROC curve can not be constructed. For instance, a non-interpretable test result can only be measured on a categorical scale. However, this result may well provide diagnostic information as there may be a relationship between non-interpretable test result and the true disease status (Begg, 1988^{B6}).

To solve this problem the likelihood ratio (LR) can be used as test parameter. The likelihood ratio is the probability of a certain test result in the presence of disease divided by the probability of the result for people without disease. If a positive test has a likelihood ratio of 5, it is 5 times more likely that this test result will be found in diseased as compared to non-diseased subjects. Thus the likelihood ratio expresses how many times more (or less) likely a test result will be found in diseased as compared to non-diseased people (Fletcher, 1988^{F1}).

The likelihood ratio has the additional advantage that should more than two test results be possible, the value of this test parameter can be calculated for each level. In the following example (Table 4.1) this will be demonstrated for stress-ECG where besides a positive or a negative test result a non-interpretable test result is also possible.

Table 4.1 Results of stress ECG for patients with (CAD+) and without (CAD-) coronary artery disease. The figures between parentheses are based on the results from our study. These data will be discussed more extensively in Chapter 7

	CAD+		CAD-	
Stress-ECG +	a	(=170)	b	(= 56)
Stress-ECG -	c	(= 58)	d	(=136)
Stress-ECG non-interpretable	e	(= 73)	f	(= 87)
	a+c+e (=301)		b+d+f (=279)	

The likelihood ratio of a positive result LR+ equals:

$$LR+ = \frac{a}{a+c+e} : \frac{b}{b+d+f} \quad (4.1)$$

or the probability of having a positive result for patients with CAD: $a/(a+c+e) = 56.5\%$, divided by the probability of having a positive result for patients without CAD: $b/(b+d+f) = 20.1\%$. Thus, the $LR+ = 2.81$.

and for the likelihood ratio of the non-interpretable test result*

$$LR_{non} = \frac{e}{a+c+e} : \frac{f}{b+d+f}$$

* For a test with a dichotomous result the LRs of the test results can be expressed in terms of sensitivity and specificity:

$$LR+ = \frac{\text{sensitivity}}{(1 - \text{specificity})} \text{ and } LR- = \frac{(1 - \text{sensitivity})}{\text{specificity}}$$

or the probability of having a non-interpretable result for patients with CAD: $e/(a+c+e) = 24.3\%$, divided by the probability of having a non-interpretable result for patients without CAD: $f/(b+d+f) = 31.2\%$. Thus, the $LR_{non} = 0.78$.

As tests are apparently not 100% foolproof one would like to estimate the probability of having disease given a certain test result. This post-test probability cannot be estimated directly from the sensitivity and the specificity or the likelihood ratio of the test alone, since it is also related to the pre-test probability of having the disease.

Applying Bayes' theorem (Bayes, 1763^{B7}), one can calculate the post-test probability, knowing the pre-test probability.

Applying the likelihood ratio, it is possible, knowing the pre-test probability, to calculate the post-test probability of having the disease given the test result and using a nomogram (Figure 4.4) (Fagan, 1975^{F2} and Sackett, 1985^{S1}).

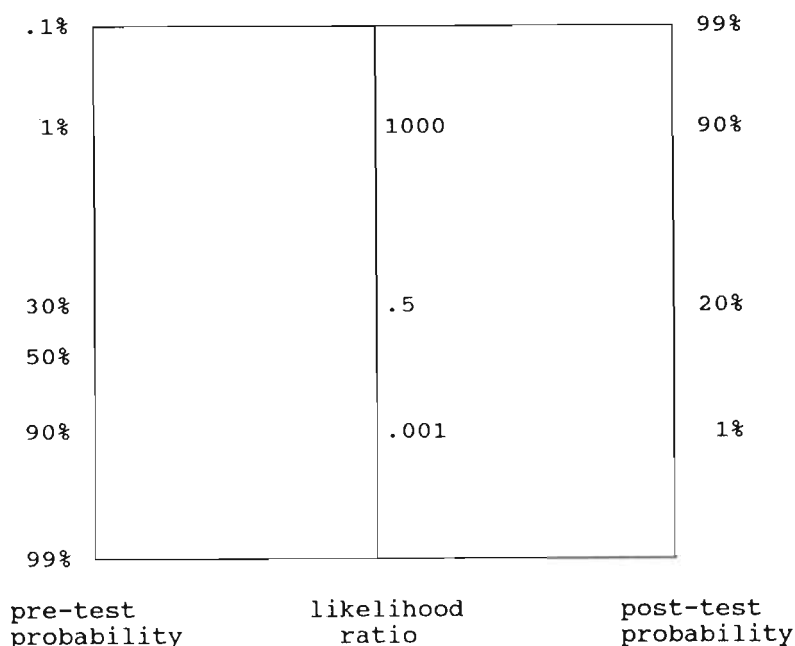


Figure 4.4 Nomogram for applying likelihood ratios

Adapted from: Nomogram for Bayes' theorem by T J Fagan

The post-test probability (right line) is found by extending a straight line from the pre-test probability value (left line) through the value for the likelihood ratio (middle line); for example the post-test probability is 20%, given a pre-test probability of 30% and an LR value of 0.5.

However, with help of the odds of having disease there is a handier way to calculate the post-test probability. (The odds of having disease is the probability of having disease divided by the probability of having no disease.)

Formula (4.1) can be rewritten as

$$LR+ = \frac{a}{b} : \frac{a+c+e}{b+d+f}$$

The first part of this expression represents the post-test disease odds, having a positive test result and the second part represents the pre-test odds.

Thus:

$$\begin{aligned} \text{LR+} &= \text{post-test odds} : \text{pre-test odds} \\ \text{or} \quad \text{post-test odds} &= \text{LR+} \times \text{pre-test odds} \end{aligned} \quad (4.2)$$

Thus, if the pre-test probability percentage is converted into a pre-test odds and multiplied by the likelihood ratio of the test result, the calculated post-test odds can be converted again to a percentage. For example, the post-test probability of a patient with a pre-test probability of 5% and a positive result of the stress-ECG can be calculated as follows.

The likelihood ratio of a positive test result equals 2.81.

A pre-test probability of 5% (or 0.05) equals a pre-test odds of

$$0.05 / 0.95 = 0.0526.$$

The post-test odds is the pre-test odds multiplied by the likelihood ratio

$$= 0.0526 \times 2.81 = 0.148.$$

A post-test odds of 0.148 is equal to a post-test probability of

$$0.148 / (1 + 0.148) = 13\%.$$

From formula 4.2 it also can be seen that if the LR of a test result equals one, the post- and pre-test probability will be equal and thus, that this test result does not discriminate between diseased and non-diseased states.

4.5 Sources of Bias

At this point it should be stressed that in determining the values of these test parameters in practice, there may be several sources of bias. This bias may partially be the cause that the values for the sensitivity and specificity (or the likelihood ratio) for the different diagnostic variables as reported in the literature have a wide range. Philbrick (1980)^{P4} reviewed the various causes leading to variations in values reported about these parameters with regard to stress ECG testing. Apart from differences in criteria defining coronary artery disease, differences in stress ECG test protocol and differences in the definition of an abnormal stress ECG, he mentioned three types of bias: (1) work-up bias (2) diagnostic review bias and (3) test review bias.

(1) Work-up bias: Sensitivity and specificity of the stress ECG designed to diagnose or evaluate coronary artery disease are traditionally determined in selected groups of subjects or patients who undergo verification of disease status by means of coronary arteriography. If the referral pattern for arteriography is biased, for example by preferential selection of patients exhibiting more severe symptoms or more extreme test results, the sensitivity and specificity observed will then be biased. Caution is needed using these biased estimates of sensitivity and specificity of the tests, especially if the tests are applied in a different population.

(2) Diagnostic review bias: the influence of the known results of the non-invasive test(s) on the arteriogram reporting.

(3) Test review bias: the influence of the result of the arteriogram on the interpretation of the non-invasive test(s).

4.6 Dependence between diagnostic variables

Bayes' theorem allows one to estimate how the outcome of a test, sign or symptom modifies the probability that a given patient does or does not have a given disease. If more tests (signs or symptoms) are conducted in sequence, the post-test probability value of the first test is the pre-test probability value of the following, second test. Independence is assumed between the different tests. At this point it suffices to say that objections have been raised against this method, often used to calculate the post-test probabilities. Various authors have indicated that the presence of dependence between tests invalidates the use of Bayes' theorem whereby independence is assumed as it may lead to inaccurate estimates of the post-test probability.

However, Sackett stated, without any argument, that this is not a problem with short sequences of two or three diagnostic tests (Sackett, 1985)^{S1}.

Apparently, this is not always the case as demonstrated in the study reported by McNeil (McNeil, 1975)^{M1}. McNeil assessed the predictive value of two tests used to diagnose reno-vascular disease, the intravenous urogram and the renogram. The gold standard for this study was the renal arteriogram. The results are demonstrated in Table 4.2.

Out of a total of 133 (=33+100) patients, 33 (=24.8%) were RVD+ and 100 (=75.2%) were RVD-. So the pretest odds were 33/100.

Table 4.2 Results of the intravenous urogram (=IVU) and the Renogram (=RG) for patients with reno-vascular disease (=RVD+) and for patients without reno-vascular disease (=RVD-)

	RVD+			RVD-		
	IVU+	IVU-	Total	IVU+	IVU-	Total
RG+	23	5	28	8	2	10
RG-	3	2	5	3	87	90
Total	26	7	33	11	89	100

From this the likelihood ratios for the positive results of each test can be calculated while assuming independence between both tests.

LR (IVU+) = (26/33)/(11/100) = 7.16

LR (RG+) = (28/33)/(10/100) = 8.48

The pretest odds being 33/100, the use of formula 4.2 allows one to determine the post-test odds for patients if both tests are positive

post-test odds = LR (IVU+) × LR (RG+) × pre-test odds

 = 7.16 × 8.48 × 33/100 = 20

Thus, the post-test probability assuming independence between the two tests, is equal to $20/(1+20) = 95\%$.

However, as can be seen from Table 4.2, in 31 patients both tests are positive of whom 23 are diseased and 8 are not. From this the post-test probability can be estimated directly and is equal to $23/31 = 74\%$.

An estimate of the post-test probability when both test are positive, but assuming independence, is higher. Apparently a positive test can lose some of its information value if known that a previous test is also positive. Therefore, dependence between both tests should be checked for and if present taken into account. One way to do this is to use a conditional LR, for example, the LR for a positive renogram, given that the IVU has been done first and was positive. Table 4.3 shows all 8 possible conditional LRs for these two tests.

Table 4.3 The LRs of the results of the renogram, conditional on the results of the IVU (= LR (RG± |IVU±)) and the LRs of the results of the IVU, conditional on the results of the renogram (=LR (IVU± |RG±))

LR (IVU+ RG+) = 1.01	LR (RG+ IVU+) = 1.22
LR (IVU- RG+) = 0.88	LR (RG- IVU+) = 0.42
LR (IVU+ RG-) = 18.0	LR (RG+ IVU-) = 32.4
LR (IVU- RG-) = 0.41	LR (RG- IVU-) = 0.28

For example, the post-test odds if each test is positive and the renogram precedes the IVU, is calculated with the following formula:

$$\text{post-test odds} = \text{LR (IVU+ |RG+)} \times \text{LR (RG+)} \times \text{pre-test odds}$$

The estimations using this formula are exactly equal to the direct estimations of the post-test probabilities.

4.7 Odds ratios

Thus, if two diagnostic variables are used to estimate the post-test probability, the above mentioned approach can be used when one checks for dependence between the diagnostic variables and if this proves to be the case conditional test parameters should be used. In our diagnostic problem the referral decision for arteriography will be based on a substantially higher number of diagnostic variables. To check if there is dependence between these variables, the LR of the results of a specific test should be determined conditionally on all the values of the other diagnostic variables. To obtain an estimate of these conditional LRs with sufficient precision, data from an enormous number of patients are necessary. To solve this problem another method for estimating disease odds can be useful. This method, a logistic regression, will be discussed in one of the following paragraphs. The use of this method implies the introduction of another test parameter, the diagnostic odds ratio (OR).

The Odds Ratio of a test result is defined as the quotient of the disease odds of that test result and the disease odds of a reference; a negative test result is usually used as this reference. Table 4.4 which is based on the results from our study on cine fluoroscopy for detecting coronary artery calcification with the arteriogram as gold standard (Chapter 9) will demonstrate this.

Table 4.4 Prevalence of coronary artery calcification in patients with (CAD+) and without (CAD-) significant coronary artery disease

	CAD+	CAD-
Fluoroscopy +	222 (a)	32 (b)
Fluoroscopy -	201 (c)	323 (d)
Total	423 (a+c)	355 (b+d)

$$OR = \frac{\text{Disease Odds with a positive result}}{\text{Disease Odds with a negative result}} = \frac{a/b}{c/d} = 11.1$$

The ORs are preferred as test parameters because the values of the ORs are not susceptible to changes in the composition of the patient population. This can also be demonstrated with help of the data from Table 4.4. As can be calculated from formula 4.1, the likelihood ratio of a positive result (LR+) equals (222/423) / (32/355)= 5.8 and for a negative result (LR-) equals (201/423) / (323/355)= 0.52. Now, if by selection mechanisms, for example a less stringent indication for coronary arteriography referral, the patient population were to consist of 50% more patients with a negative test result (c= 302 and d= 485), the LR+ would be 6.8 and the LR- would be 0.61. However, the odds ratio in this case remains 11.1 (=(222/302)/(32/485)) and in contrast with the LRs does not change. Figure 4.5 may serve to demonstrate this phenomenon as well as how the discriminative value of a test is described by the OR and the LR as test parameters. This figure illustrates that the estimate of the disease odds increases with a positive fluoroscopy and decreases with a negative result. The OR is the quotient of both post-test Odds and the discriminating value of the test will be high if the difference between both Odds is great and consequently the OR will be high. The LR is the quotient between the post-test Odds and pre-test Odds and the value of this parameter will be determined by the difference of the levels of the post-test and the pre-test Odds. It is assumed that the post-test Odds are non-variable measures for the disease probability and therefore the OR is a non-variable parameter. However, the pre-test Odds are also dependent on the numbers of the patients with positive and negative results in the study population and can theoretically have any value between the two post-test Odds. Thus, the LR is variable and any change in the study population can influence the value of the LRs.

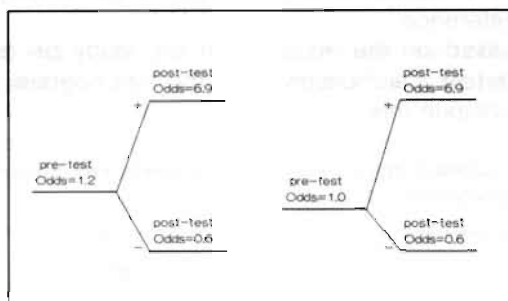


Figure 4.5 Pre- and post-test Odds of fluoroscopy for two populations differing in pre-test Odds, demonstrating the difference between the LR and OR as test parameters.

$$OR = \frac{\text{posttest Odds} +}{\text{posttest Odds} -} = 11.1$$

$$LR+ = \frac{\text{posttest Odds} +}{\text{pretest Odds}} = 5.8 \text{ or } 6.8$$

$$LR- = \frac{\text{posttest Odds} -}{\text{pretest Odds}} = 0.52 \text{ or } 0.61$$

4.8 Collinearity and interaction

The main reason for using the OR as test parameter in diagnostic probability analysis is the possibility of taking account of the complex influence of the mutual dependence between the diagnostic variables on the estimations of post-test probabilities. The influence of this dependence on the probability estimations can be described by two mechanisms: collinearity and interaction. The effect of collinearity and interaction are demonstrated in the following somewhat exaggerated examples.

A smart doctor who wanted to know if his patients had cardiovascular disease asked them to bare their legs, and he evaluated carefully the extent of leg hair. He divided the patients in two groups, one having 'much hair' (H+), the other having 'no hair' (H-). All these patients underwent coronary arteriography so he was able to construct a table in which he could correlate the amount of hair with the presence (CAD+) or absence (CAD-) of disease (Table 4.5).

Table 4.5 Result of test H and coronary artery disease (CAD)

	CAD+	CAD-		
H+	52	20	72	sensitivity = 52/72
H-	20	52	72	specificity = 52/72
	72	72	144	discriminative value of this test OR = 6.76

He discovered that if a patient had 'much hair', the chance of having coronary artery disease was also very high. The test appeared to have a good discriminative value, having an Odds Ratio of 6.76. However, one of his colleagues proposed

analysing his findings further, stratifying for sex. The results are demonstrated in Table 4.6.

Table 4.6 Results of test H and coronary artery disease stratified for gender

men				women			
	CAD+	CAD-			CAD+	CAD-	
H+	50	10	60	H+	2	10	12
H-	10	2	12	H-	10	50	60
	60	12	72		12	60	72

Sensitivity = 50/60
Specificity = 2/12
discriminative value of this test
OR = 1

Sensitivity = 2/12
Specificity = 50/60
OR = 1

Apparently, sex influences the function of the test, and the presence of hair does not discriminate any more between diseased and non-diseased. This is not surprising, since men have more hair on their legs than women, and men have, until now, more often coronary artery disease. Apparently hair on the legs and sex are interrelated, and when sex is used as a diagnostic variable, the presence of hair on the legs does not contribute any more in predicting the presence or absence of coronary artery disease. In other words, once a male group has been separated by stratification according to gender, hair does not have any predictive value. It is therefore important to take into account underlying relationships of the various diagnostic variables. Collinearity is present if the association between the test result and the disease occurs (at least partly) indirectly through another variable (Fig. 4.6). In this example this variable appears to be sex. The predictive value of gender was falsely attributed or confounded with the predictive value of the presence of hair on the legs.

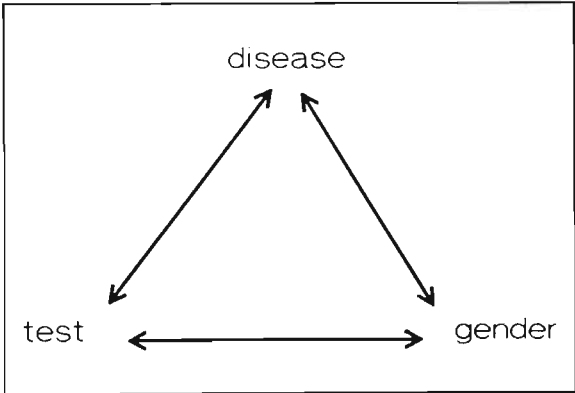


Figure 4.6 Relationship between a test, a diagnostic variable and disease in the presence of collinearity

It can also be seen from this example that an approach using the OR demonstrates this effect more clearly than using the traditional concepts of sensitivity and specificity. We now only have to compare the OR=6.67 of males and females

together, with the OR of males (OR=1) and females (OR=1) separately, which is easier to do than to compare all values found for the sensitivity and specificity. Another reason for using OR instead of sensitivity and specificity will become apparent after we have introduced the logistic regression model. An additional mechanism called interaction can occur when the strength of the relationship between the test and the disease differs according to the level of a third variable (Figure 4.7).

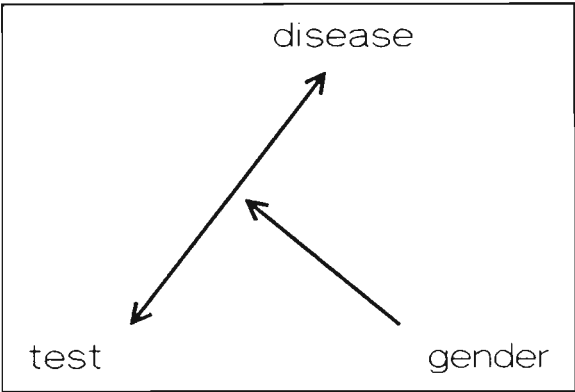


Figure 4.7 Strength of relationship between test and coronary artery disease depends on the level of a third variable, demonstrating the presence of interaction

If this type of interference had been present in our doctor's observations, the results of his observations could have been as follows (Table 4.7).

Table 4.7 Results of test H and coronary artery disease stratified for gender, with interaction present

men				women			
	CAD+	CAD-			CAD+	CAD-	
H+	55	10	65	H+	2	10	12
H-	5	2	7	H-	10	50	60
	60	12	72		12	60	72
discriminative value of this test							
OR = 2.2				OR = 1			

In this situation, the relationship between having hairy legs and having coronary artery disease appears to be reinforced if the patient is a male. In the previous example, we assumed that 50 out of 60 male patients with coronary artery disease had much hair; in this example there are 55 out of 60 males with coronary artery disease who are positive on the criterion of much hair. So, the OR appears to be (somewhat) higher for men (2.2) than for women (1.0). Thus interaction is present and this has to be accounted for. In a situation with no interaction, the OR will be the same for each stratum within a category. In a situation in which interaction is present, the OR is different for

each stratum of a certain category. This is not a surprise as both situations give a different description of the underlying relationship among the variables involved. Another way to uncover the presence of interaction is demonstrated in the following approach. Consider two dichotomous (0,1) variables, X_1 and X_2 . ORs can be defined comparing the odds for any cell in the two-way table with the odds for any other cell. The reference cell is usually selected in which both variables have the lowest value, in our case comparing all other possibilities with the situation in which both variables are negative (thus $X_1 = 0$ and $X_2 = 0$).

	$X_1 = 0$	$X_1 = 1$
$X_2 = 1$	odds ₁₁ = a	odds ₁₀ = b
$X_2 = 0$	odds ₀₁ = c	odds ₀₀ = d

The OR is thus defined as OR_{11} , OR_{10} and OR_{01} where

$$OR_{11} = Odds_{11} / Odds_{00} = a/d$$

$$OR_{10} = Odds_{10} / Odds_{00} = b/d$$

$$OR_{01} = Odds_{01} / Odds_{00} = c/d$$

How can the presence of interaction now be detected? In order to detect the presence of interaction two situations will be compared: in the first situation the effect is considered where both variables have the value 1 at the same moment (situation a); in the other situation the effect of variable X_1 having the value 1 and acting alone (situation c) is multiplied by the effect of variable X_2 having the value 1 and also acting alone (situation b). Having defined the ORs for these situations, the following equation allows assessment of multiplicative interaction:

$$OR_{11} = OR_{10} \times OR_{01}$$

This indicates that the effect of both variables acting together is the same as the multiplicative result of the effect of each variable separately. If this expression is satisfied there is no interaction on a multiplicative scale. On the contrary, if the expression is not satisfied, interaction on a multiplicative scale is deemed to be present.

Consider once again our examples (Table 4.6 & 4.7). In the first example, there was no interaction.

		CAD+	CAD-	
men	H+	odds = 5	odds = 5	$OR_{11} = OR_{10} \times OR_{01}$
	H-	odds = 0.2	odds = 0.2	

In the example used to explain the presence of interaction, the OR_{11} is not equal to $OR_{10} \times OR_{01}$.

		CAD+	CAD-	
men	H+	odds = 11	odds = 5	$OR_{11} \neq OR_{10} \times OR_{01}$
	H-	odds = 0.09	odds = 0.2	

Therefore, one possibility of showing the presence of interaction is to compare the OR between strata; the other way is to use the rule: OR_{11} equals $OR_{10} \times OR_{01}$ means no interaction. Thus, the diagnostic variables, in this example the gender of the patient and the hair on the legs, may be interrelated, a condition which may not be neglected in analysing our problem. Therefore, a method is needed to disentangle the interrelationships between diagnostic variables (or predictors) and disease status. As shown above, these interrelationships can be described by collinearity and interaction and thus a correct probability analysis has to take account of these two mechanisms. A method that is increasingly used for this kind of analysis is logistic regression. As will be elaborated in the next section, the logistic regression model expresses the discriminative value of a (diagnostic) variable with help of the (diagnostic) odds ratio (OR). A computer program using this model can be used to provide estimates for the values of the ORs for all the results of the diagnostic variables.

4.9 Logistic Regression as method for estimating an OR

4.9.1 Introduction

The probability of a patient having disease can be predicted by various diagnostic variables, which can be measured by several tests, including history-taking. Often some of these variables are more or less mutually dependent, therefore providing less or even no additional information. Thus, a model is needed to untangle the mutual dependence in the complex relationships between these variables and the probability of disease in order to assess the independent contribution of each variable. As shown in the previous paragraphs, for this analysis the OR is the test parameter of choice. In fact the analysis (with help of the logistic regression model) will provide an estimate of the disease odds dependent on the diagnostic characteristics.

Suppose that for a given patient information on a number of diagnostic variables X_1, X_2, \dots, X_k is collected. With help of this information the disease probability can be estimated and will be expressed as: $P(D^+ | X_1, X_2, \dots, X_k)$, meaning the probability P of having the disease status D^+ given the values of the diagnostic variables X_1 through X_k , which are specific for each patient.

The disease Odds are defined as

$$\text{Odds}(D^+ | X_1, X_2, \dots, X_k) = P(D^+ | X_1, X_2, \dots, X_k) / (1 - P(D^+ | X_1, X_2, \dots, X_k))$$

In the following text $P(D^+ | X_1, X_2, \dots, X_k)$ will be abbreviated as $P(\mathbf{X})$ and $\text{Odds}(D^+ | X_1, X_2, \dots, X_k)$ as $\text{Odds}(\mathbf{X})$. Thus,

$$\text{Odds}(\mathbf{X}) = P(\mathbf{X}) / (1 - P(\mathbf{X}))$$

Logistic regression is a mathematical modelling approach that can be used to describe this relationship between several variables and the disease status with help of these $\text{Odds}(\mathbf{X})$. There are other modelling approaches, but for diagnostic

analyses logistic regression has several advantages as will be shown in the following. The basic concepts of this method will be discussed first.

4.9.2 The logistic function

Logistic regression is a popular mathematical modelling procedure used in the analysis of data. In epidemiological research a typical question is to establish the relationship of one or more independent variables to a disease status. Here this method will be used to analyse how diagnostic variables such as test results, symptomatology, age and gender are associated with the presence of coronary artery disease.

The chance that a patient has a disease is expressed on a scale ranging from 0 to 1 (Figure 4.8). The condition of the patient can be such that the chance of having disease is negligible; thus, a constellation of values of variables does exist in which this chance is practically 0. In the same way, there are many possible combinations of values of variables, from which it is practically certain that a patient does have the disease. Between these two points there is a range where the probability of having disease changes from 0 to 1.

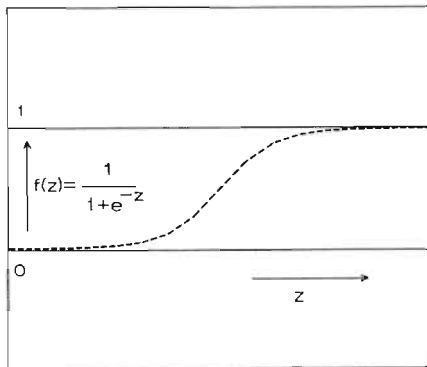


Figure 4.8 Probability of having disease ranging from 0 to 1, as a function of z which represents the cumulative value of several variables

For many problems this alteration of the probability can be described by an S curve. In the middle portion of this S-shaped curve a physician usually has to make a decision, whereas in the extremities of the curve, the disease status of the patient is clear.

As indicated in the figure this S-shaped curve can be described by a logistic function. This function has the advantageous property that it ranges between 0 and 1 and therefore the probability estimate, by definition, will never be below 0 or above 1.

This logistic function is defined as:

$$f(z) = 1 / [1 + e^{-z}]$$

where $f(z)$ represents the probability of having the disease and z is the linear expression of all diagnostic variables involved:

$$z = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

where $X_1 \dots X_k$ are the values of the (diagnostic) variables of interest belonging to patient X and α, β are constant terms representing unknown coefficients, which have to be estimated.

When z is substituted into the equation the logistic model can be written in the following way:

$$P(D^+ | X_1, X_2 \dots X_k) = P(X) = 1 / [1 + e^{-(\alpha + \sum \beta_i X_i)}] \quad (4.9.2.1)$$

For all of the diseased as well as the non-diseased patients of the study population the combination of X_1 through X_k is known, and the coefficients α and β_i ($i=1,2,\dots,k$) can be estimated by maximizing the likelihood function. This statistical procedure is similar to the familiar linear regression. This can be shown by the following transformation (logit transformation).

$$\text{If } P(X) = 1 / [1 + e^{-(\alpha + \sum \beta_i X_i)}]$$

$$\text{then Odds}(X) = P(X) / (1-P(X)) = e^{\alpha + \sum \beta_i X_i} \quad (4.9.2.2)$$

$$\text{and } \log[\text{Odds}(X)] = \alpha + \sum \beta_i X_i$$

with $\log[\text{Odds}(X)]$ the natural logarithm of $\text{Odds}(X)$. This is the wellknown expression of the linear regression model with the $\log(\text{Odds}(X))$ as dependent variable. The procedure generates a logistic function that fits the empirical data best. After this has been done, the model can be used to calculate the probability of having disease for any new patient for whom the values for the variables are known.

4.9.3 Logistic function and Odds Ratio

The discriminating value of a test can be expressed by an Odds Ratio. As shown in the example of Tables 4.5 and 4.6 the OR has to be determined stratified for all other diagnostic variables, i.e. these other diagnostic variables have all the same value for this 2x2-table. In this example these ORs could be calculated readily. However, if more diagnostic variables are involved a (logistic) model is needed to estimate the ORs.

Thus, although it is not necessary to fit a logistic model to get an OR estimate, the examples in tables 4.5, 4.6 and 4.7 will be used to demonstrate how the OR can be determined with help of the logistic regression model, taking collinearity and interaction between the diagnostic variables in account.

First the problem will be formulated in a notation in conformity with formula (4.9.2.1)

The dependent variable is coronary artery disease with two outcome possibilities: D^+ and D^- . Furthermore, there are two diagnostic variables:

X_{sex} with the levels 1 for males and 0 for females.

X_{hair} with the levels 1 if the hair test is positive and 0 if the test is negative.

The probability of having disease given the result of the hair test and gender is

$$P(D^+ | X_{\text{hair}}, X_{\text{sex}}) = P(\mathbf{X}) = 1/[1 + \exp(-(\alpha + \beta_{\text{hair}} \cdot X_{\text{hair}} + \beta_{\text{sex}} \cdot X_{\text{sex}}))] \quad (4.9.3.1)$$

and from (4.9.2.2) it follows that

$$\text{Odds}(\mathbf{X}) = \exp(\alpha + \beta_{\text{hair}} \cdot X_{\text{hair}} + \beta_{\text{sex}} \cdot X_{\text{sex}}) \quad (4.9.3.2)$$

With help of this formula the Odds(\mathbf{X}) can be determined for all combinations of different values of the diagnostic variables:

females

$$\text{hair test negative: Odds}_{\varnothing, h^-} = \exp(\alpha) \quad (4.9.3.3)$$

$$\text{hair test positive: Odds}_{\varnothing, h^+} = \exp(\alpha + \beta_{\text{hair}}) \quad (4.9.3.4)$$

males

$$\text{hair test negative: Odds}_{\delta, h^-} = \exp(\alpha + \beta_{\text{sex}}) \quad (4.9.3.5)$$

$$\text{hair test positive: Odds}_{\delta, h^+} = \exp(\alpha + \beta_{\text{hair}} + \beta_{\text{sex}}) \quad (4.9.3.6)$$

By dividing (4.9.3.4) by (4.9.3.3) and (4.9.3.6) by (4.9.3.5) the OR of the hair test for females and for males respectively can be determined and this OR_{hair} is in both cases equal to $\exp(\beta_{\text{hair}})$.

In the same way the diagnostic OR of gender for patients with negative or positive results of the hair test can be deduced by dividing (4.9.3.5) by (4.9.3.3) and (4.9.3.6) by (4.9.3.4) respectively. The OR_{sex} is for both results of the hair test $\exp(\beta_{\text{sex}})$. Obviously there is a simple relation between the OR and coefficients of the logistic function as the OR of a diagnostic variable is equal to the exponent of the corresponding coefficient β . Furthermore $\exp(\alpha)$ is equal to the disease Odds of the reference category, i.e. the category of patients for whom all the X's are zero. In this example this reference category consists of females with a negative hair test result.

For this example, logistic regression would generate the following values for the coefficients.

$$\alpha = -1.61 \quad [\text{Odds}_{\varnothing, h^-} = 10/50 = 0.2 = e^{-1.61}]$$

$$\beta_{\text{hair}} = 0 \quad [OR_{\text{hair}} = 1 = e^0]$$

$$\beta_{\text{sex}} = +3.22 \quad [OR_{\text{sex}} = (10 \times 50)/(2 \times 10) = 25 = e^{3.22}]$$

Summarizing, by estimating ORs starting from the Odds of all possible combinations of the diagnostic variables, the logistic regression model takes account of collinearity.

The numbers in the cells of Table 4.5 and 4.6 have been chosen in such a way that the OR_{hair} is the same for males and females and that the OR_{sex} is the same

for a positive and a negative test result. Stated differently, there is no interaction and the model as described in (4.9.3.1) estimates the probabilities exactly.

$$\text{As Odds}_{\delta,h+} = \exp(\alpha + \beta_{\text{hair}} + \beta_{\text{sex}}) = \exp(\alpha)^* \times \exp(\beta_{\text{hair}}) \times \exp(\beta_{\text{sex}})$$

and substituting the expression for the ORs and the Odds of the reference category,
the following holds if there is no interaction

$$\text{Odds}_{\delta,h+} = \text{Odds}_{\varnothing,h-} \times \text{OR}_{\text{hair}} \times \text{OR}_{\text{sex}} \quad (4.9.3.7)$$

In the example of Table 4.7 the OR of the hair test is equal to 2.2 for males and to 1 for females and OR_{sex} is equal to 27.5 for a positive hair test result and 12.5 for a negative result and (4.9.3.7) does not hold as can easily be shown.

$$\text{Odds}_{\delta,h+} \neq \text{Odds}_{\varnothing,h-} \times \text{OR}_{\text{hair}} \times \text{OR}_{\text{sex}} \quad (4.9.3.8a)$$

$$\text{as } 55/10 \neq 10/50 \times 1 \times 12.5 \quad (4.9.3.8b)$$

If the logistic regression model of the previous example were also to be used for this example of Table 4.7 the computer would estimate an average OR_{hair} for males and females and an average OR_{sex} for both test results, ignoring a (possible) interaction between gender and the hair test.

If the model has to take account of interaction an additional (interaction) term has to be put into it and the general formula for the Odds(X) will become

$$\text{Odds(X)} = \exp(\alpha + \beta_{\text{hair}} \cdot X_{\text{hair}} + \beta_{\text{sex}} \cdot X_{\text{sex}} + \beta_{\text{hair,sex}} \cdot X_{\text{hair}} \cdot X_{\text{sex}}) \quad (4.9.3.9)$$

The formulae (4.9.3.3) through (4.9.3.5) remain the same and the formula for $\text{Odds}_{\delta,h+}$ becomes

$$\begin{aligned} \text{Odds}_{\delta,h+} &= \exp(\alpha + \beta_{\text{hair}} + \beta_{\text{sex}} + \beta_{\text{hair,sex}}) \\ &= \exp(\alpha) \times \exp(\beta_{\text{hair}}) \times \exp(\beta_{\text{sex}}) \times \exp(\beta_{\text{hair,sex}}) \end{aligned} \quad (4.9.3.10)$$

And substituting the expressions of the ORs and the reference Odds:

$$\text{Odds}_{\delta,h+} = \text{Odds}_{\varnothing,h-} \times \text{OR}_{\text{hair}} \times \text{OR}_{\text{sex}} \times \exp(\beta_{\text{hair,sex}}) \quad (4.9.3.11)$$

From (4.9.3.8b) it can be calculated that the factor $\exp(\beta_{\text{hair,sex}})$ must have the value 2.2 to rectify the inequality and to ensure that (4.9.3.10) holds.

Summarizing, by putting this kind of additional terms in the logistic model, interaction between the diagnostic variables can be accounted for.

* $\exp(\alpha) = e^{\alpha}$

In this simple example the coefficients could also be calculated directly. If more diagnostic variables are involved, a computer program is required to estimate the coefficients (the α and the various β coefficients) of the logistic model.

However, first a choice has to be made whether or not to include interaction terms in the model. If there are more than two diagnostic variables, higher order interaction terms might also be included.

This decision depends on whether these terms can be interpreted and whether the fit of the data has improved significantly. For our computations the BMDP (Biomedical DATA package) computer program has been used which provides statistics to assess this fit of the data. In our model only those first-order terms were used which could be interpreted.

5

Determination of $P_{\text{threshold}}$ indicating when an arteriogram should be performed

5.1 $P_{\text{threshold}}$

To construct a diagnostic protocol for patients with symptoms suggesting coronary artery disease, one first has to decide on a criterion which patient should have an arteriogram, or more precisely for which combination of diagnostic variables arteriography should be advised. Whether patients with certain diagnostic characteristics need referral for arteriography is ultimately determined by the average clinical improvement of the patients as the result of such a referral. Not only the clinical variables but also the results of various non-invasive tests will be the yardstick to decide which patient will have an arteriogram. It will be assumed that the probability of having coronary artery disease, based on these diagnostic variables will determine the referral decision.

In this chapter it will be shown that the criterion for this referral decision can be formulated with help of the $P_{\text{threshold}}$. This term, $P_{\text{threshold}}$, has been introduced in Chapter 2. A patient with a probability of having coronary artery disease higher than the $P_{\text{threshold}}$ will have an arteriogram, while a patient with a lower value will not.

5.2 Determination of $P_{\text{threshold}}$ by means of a medical decision analysis

The classical method for establishing this $P_{\text{threshold}}$ is by way of a medical decision analysis or a cost-effectiveness analysis in which patients who have been referred are compared with patients who have not been referred. The clinical outcome of both groups of patients is estimated by adding the utilities of the different results of the various patients, weighted with the probability of such results. The average difference between the two groups, eventually also accounting for cost aspects, determines the value of the arteriography. The problem of the referral decision can be schematically described by the diagram shown in Figure 5.1.

The averaged outcome of performing an arteriogram for N patients can be presented as:

$$P \times N \times (D + B_{\text{therapy}}) + (1 - P) \times N \times (H - L_{\text{arteriography}}) \quad (5.1)$$

The averaged outcome of not performing an arteriogram for N patients can be presented as:

$$P \times N \times D + (1 - P) \times N \times H \quad (5.2)$$

The benefit of a referral for arteriography (and its consequences) over non-referral for N patients is the difference between expression 5.1 and 5.2:

$$\begin{aligned}
 &= P \times N \times (D + B_{\text{therapy}}) + (1-P) \times N \times (H - L_{\text{arteriography}}) - P \times N \times D - (1-P) \times N \times H \\
 &= P \times N \times B_{\text{therapy}} - (1-P) \times N \times L_{\text{arteriography}} \quad (5.3)
 \end{aligned}$$

It is clear that if this value is negative no arteriogram will be performed. When the probability of having disease equals $P_{\text{threshold}}$ there is no preference whether or not to perform an arteriography and thus the following equation holds:

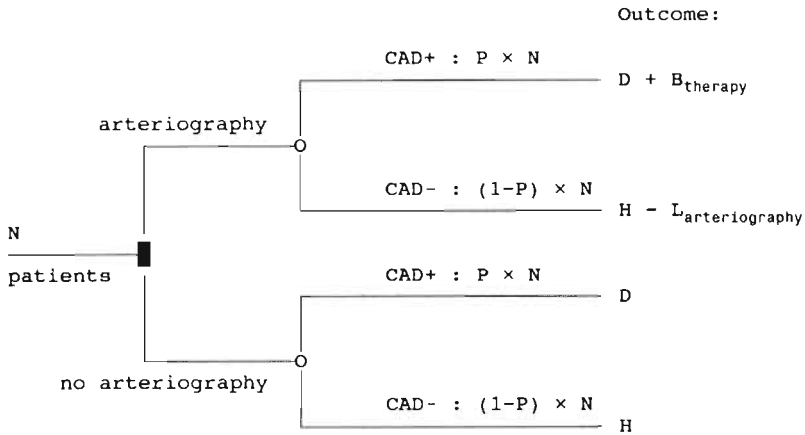


Figure 5.1 Outcome arteriography/no arteriography, with P as the chance of having disease

Where:

N = the number of patients; P = post-test probability of having coronary artery disease; thus, $P \times N$ is the number of patients with disease and $(1-P) \times N$ is the number of patients without disease; H = the quality of life belonging to a healthy status; D = the quality of life belonging to a diseased status; B_{therapy} = the net profit of an arteriogram; this is defined as the improvement of the quality of life or the increase in life expectancy and life quality thanks to therapy, less the risks for the patient and the cost of this therapy and minus the loss due to the arteriography* ($L_{\text{arteriography}}$) as described below. This benefit will be expressed in B_{therapy} units; $L_{\text{arteriography}}$ = the loss due to performing an arteriogram on a healthy person; this is defined as certain negative effects such as the risk of death or morbidity due to the arteriogram, the inconveniences, and the cost of the arteriogram. This loss will be expressed in $L_{\text{arteriography}}$ units; CAD+/- = whether or not coronary artery disease is present.

* The risk of complications of arteriography is different in diseased and non-diseased patients and the value of $L_{\text{arteriography}}$ will be different in the two categories. For the diseased patients this is accounted for in the term B_{therapy} .

$$P_{\text{threshold}} \times N \times B_{\text{therapy}} = (1 - P_{\text{threshold}}) \times N \times n L_{\text{arteriography}} \quad (5.4)$$

and from this

$$P_{\text{threshold}} = L_{\text{arteriography}} / (L_{\text{arteriography}} + B_{\text{therapy}}) \quad (5.5)$$

Although feasible, it is in practice difficult to obtain relevant data on utilities of the various results as well as the probabilities. The possible outcomes of decisions whether or not to perform an arteriogram, differences in treatment varying with the lesions found, etc., have all to be expressed numerically. Collecting this kind of data is very cumbersome, and furthermore most of data are not complete, are often refutable and may well change in the course of time.

The benefit of coronary artery bypass grafting at present is different from the situation in 1970 when the last large study of coronary artery bypass grafting, the coronary artery surgery study (CASS), was performed. A recent study reported by Neutze (1987)^{N4} reported a higher reduction in mortality from coronary artery disease by cardiac surgery than previous studies. This is indicative of the continuous change of survival rate data and points to the difficulty in obtaining the reliable data needed.

Furthermore, the effect of the recent, widespread application of internal mammary artery bypass grafting has still not been fully assessed and the same applies to the Percutaneous Transluminal Coronary Angioplasty (PTCA) procedure (balloon coronary angioplasty).

Moreover, development of new drugs and procedures and refinements in established techniques cause changes and hinder evaluation of outcome in comparing medical and surgical therapies in coronary artery disease (Pryor, 1987^{P5}). Killip and Cameron concur that it is very difficult to quantify the outcome of the different possibilities when dealing with patients suspected of having coronary artery disease (Killip, 1988^{K4} and Cameron, 1990^{C1}).

In conclusion, although a decision analysis for this problem is theoretically possible, it is in practice difficult to obtain relevant data on the various outcome variables as well as the probabilities. Inclusion of costs in the analyses would complicate the evaluation further.

For the decision analysis, the missing information could be filled in by the best possible suppositions, thus enabling the formulation of a referral advice for the different categories of patients. However, if this advice were to be far beyond the present clinical practice, the results of an analysis would not be believed by clinicians and would be deemed unacceptable. As a consequence these suppositions would be adapted to obtain a more acceptable result and this adaptation process would go on until the referral advices are similar to the present work-up. This process could be short-circuited by replacing this kind of decision analysis

* Dividing (5.4) by N gives:

$$\begin{aligned} P_{\text{threshold}} \times B_{\text{therapy}} &= (1 - P_{\text{threshold}}) \times L_{\text{arteriography}} \\ &= L_{\text{arteriography}} - P_{\text{threshold}} \times L_{\text{arteriography}} \\ P_{\text{threshold}} \times B_{\text{therapy}} + P_{\text{threshold}} \times L_{\text{arteriography}} &= L_{\text{arteriography}} \\ P_{\text{threshold}} \times (B_{\text{therapy}} + L_{\text{arteriography}}) &= L_{\text{arteriography}} \\ P_{\text{threshold}} &= L_{\text{arteriography}} / (L_{\text{arteriography}} + B_{\text{therapy}}) \end{aligned}$$

by a consensus meeting of all involved (medical specialists, hospital board, insurance companies, and organisations representing patients) to formulate the indications for arteriography.

5.3 Determination of $P_{\text{threshold}}$ based on the budget

Although determination of $P_{\text{threshold}}$ by means of a medical decision analysis would be possible, another method has been chosen to establish the $P_{\text{threshold}}$. Here the limitation of funds due to the compulsory budget of the hospital has been used to determine the $P_{\text{threshold}}$. This approach will be elaborated in the following.

In the De Wever Hospital in Heerlen, the number of coronary arteriograms for the years 1987 and 1988 was frozen at 500 per year, decided by the board of directors in order to stem rising costs, 500 being the average number of coronary arteriograms performed in the preceding years. To select the patients for an arteriogram, knowledge of the population visiting the cardiologists is needed. Details regarding numbers and characteristics may be drawn from the population referred for arteriography during the period 1983 - 1987. This population, consisting of 778 patients, will serve as a model.

On the basis of the records of the above-mentioned cardiologists it is expected that of the 500 arteriograms which may be performed, 275 would have to be used to assess patients with a known myocardial infarction, valvular disease, etc. Furthermore, there will be 30 patients with unstable angina, who will have an arteriogram without preliminary non-invasive testing. The protocol to be developed will not be applicable to these patients. The remaining 195 arteriograms would need to be divided among the remaining patients requiring an arteriogram to confirm or to rule out coronary artery disease. To select those patients who will have the highest average clinical improvement as a result of arteriography, an estimation of the disease probability on the basis of the diagnostic characteristics is required. And for this estimation one needs information about the relation between diagnostic variables and the disease status proven by arteriography. This information may be drawn from the records of patients with chest pain referred to the cardiologists during the 1983 - 1987 period. Because the disease status has to be known, only those patients in whom arteriography has been performed can be used. Patients who have not been referred for arteriography are not included and their data are not used in the analysis. Consequently the protocol to be developed can only be applied to patients who would be referred if the referral pattern was unchanged. Furthermore, it could be expected that the new protocol would also rule out arteriography in those patients now treated without an arteriogram.

It is reasonable to expect that the new protocol has to be more selective. The number of patients without significant coronary artery disease appeared relatively high during the years 1983-1987. The prevalence of coronary artery disease in the referred patients was 54% which was rather low as compared with published findings (Phibbs, 1988^{P6}). This figure is arguable in view of the current waiting list and the expected increase in demand due to the aging of the population. It is also likely that the need for diagnostic arteriograms will increase with the development of new techniques such as a PTCA procedure enabling intervention in more patients. As a consequence of these considerations the protocol to be developed should be more selective as to whether or not to refer.

In view of the above-mentioned assumptions and data, the number of patients referred for arteriography is estimated at 250 each year. As only 195 may be performed, about 55 patients (22%) need to be selected out by the protocol not to receive an arteriogram. Choosing the 55 patients with the lowest chance of having coronary artery disease may be the best selection criterion.

If all the non-invasive tests were performed in those 250 patients and thus the maximum information was available, those 250 could be ranked in order of chance of having coronary artery disease. Only the 195 patients with the highest chance would then have an arteriogram. A list assembled in such a way provides a chance value above which arteriography is still performed and below which the patient is denied arteriography. The chance at which the decision changes must be equal to the $P_{\text{threshold}}$ as derived in (5.5).

Two problems arise in the production of such a list. First, the logistic regression model, as discussed in Chapter 4, only provides an estimation of the probability of disease for all of the 288 combinations of diagnostic characteristics. However to compile a list of the 778 patients, ranked in order of chance of having coronary artery disease, the number of patients for each of these 288 categories is also needed. As described in the appendix of this chapter a hierarchical log-linear model was used for this purpose. Secondly, not all the non-invasive tests have been done for all the patients. For a patient in whom a test had not been performed, conditionally on the disease status, an estimate of the probabilities of all possible test results is required. Also in the appendix of this chapter the method is described how these probabilities are estimated on the base of all known other diagnostic data of that patient. The resulting list of 778 patients, ranked in order of chance of having coronary artery disease, is demonstrated in Table 5.1.

As deduced above, the advice would be to withhold coronary arteriography in about 22% of the patients. For the 778 patients who were collected during a period of 50 months, it would mean that 171 of them would not have an arteriogram. According to the list compiled in Table 5.1 this results in a decision criterion to defer arteriography if the post-test probability is less than 15 %. Thus, the exclusion of 55 of 250 patients on an annual basis results in a $P_{\text{threshold}}$ of 15%.

Establishing 15% as the $P_{\text{threshold}}$ seems reasonable; a post-test probability of < 15% seems justified as a level to withhold arteriography. Although not explicitly stated, the cardiologist accepts a certain percentage of uncertainty when he denies a patient arteriography (Hung, 1985^{H3}). Steingart (1990)^{S4} ranks a post-test probability < 15% as low.

5.4 Conclusion

Assuming that in the future the composition of the population referred for arteriography will not change dramatically compared with recent years, the decision of the board of directors to freeze the number of arteriograms at 500 per year implies that patients referred for arteriography will undergo an arteriogram if the post-test probability is $\geq 15\%$, while patients with a post-test probability < 15% will be refused an arteriogram. A shortage of coronary arteriography facilities for diagnostic purposes could persist in the future even if this sharper selection is imposed, based on the proposed $P_{\text{threshold}}$ of 15%. A request for a higher budget, for more coronary arteriograms, could then be considered and would be justified, based on the data presented here.

Table 5.1 Cumulative distribution of post-test probability for all possible combinations of diagnostic variables

Post-test prob. %	Cumulative number of patients*	Sex	Age	Sympt	T	S	C
.992	25.51	M	65	+++	+	++	+
.992	44.83	M	55	+++	+	++	+
.990	47.97	M	65	+++	-	++	+
.990	50.53	M	55	+++	-	++	+
.985	60.57	M	45	+++	+	++	+
.982	64.90	F	65	+++	+	++	+
.981	78.16	F	65	+++	+	++	+
.981	79.42	M	45	+++	-	++	+
.977	82.66	F	65	+++	-	++	+
.977	83.72	F	55	+++	-	++	+
.974	92.39	M	65	+++	+	+	+
.974	102.76	M	55	+++	+	+	+
.968	104.04	M	55	+++	-	+	+
.967	105.11	M	65	+++	-	+	+
.964	106.11	F	45	+++	+	++	+
.955	106.36	F	45	+++	-	++	+
.949	110.96	M	45	+++	+	+	+
.947	128.71	M	65	+++	+	N	+
.947	143.07	M	55	+++	+	N	+
.940	144.79	F	55	+++	+	+	+
.939	148.12	F	65	+++	+	+	+
.937	148.70	M	45	+++	-	+	+
.936	159.65	M	65	+++	+	++	-
.936	182.82	M	55	+++	+	++	-
.934	185.04	M	65	+++	-	N	+
.934	186.83	M	55	+++	-	N	+
.933	198.70	M	55	+++	+	-	+
.932	204.75	M	65	+++	+	-	+
.927	206.10	M	65	++	+	++	+
.926	206.53	F	55	+++	-	+	+
.925	207.42	M	55	++	+	++	+
.925	208.24	F	55	+++	-	+	+
.921	211.14	M	55	+++	-	++	-
.920	212.51	M	65	+++	-	++	-
.917	213.99	M	55	+++	-	-	+
.916	214.75	M	65	+++	-	-	+
.910	215.31	M	65	++	-	++	+
.908	215.69	M	55	++	-	++	+
.899	217.66	M	45	+++	+	N	+

* For example, the number 60.57, in the second column, in the fifth row, means that 60.57 persons have a probability of $\geq 98.5\%$ of having coronary artery disease.

Sex M/F = male/female; age 45/55/65 = <50 / 50-59/ ≥ 60 ; symptoms = +++ typical ap/ ++ atypical ap / + n.c.p./asymptomatic; T +/- = +/- thallium scintigram; S -/+ / ++ / N = negative stress ECG / + positive stress ECG with 1-2 mm ST segment depression / ++ positive stress ECG with ≥ 2 mm ST segment depression / N non-diagnostic test result; +/- C = +/- cine fluoroscopy.

continuation (1) Table 5.1

Post-test prob. %	Cumulative number of patients	Sex	Age	Sympt	T	S	C
.885	218.00	F	45	+++	+	+	+
.883	221.29	F	55	+++	+	N	+
.881	230.76	F	65	+++	+	N	+
.879	247.46	M	45	+++	+	++	-
.876	247.71	M	45	+++	-	N	+
.873	253.77	M	45	+++	+	-	+
.862	254.56	M	45	++	+	++	+
.860	254.64	F	45	+++	-	+	+
.858	254.93	M	65	+	+	++	+
.856	255.75	F	55	+++	-	N	+
.856	256.10	M	55	+	+	++	+
.854	258.47	F	65	+++	-	N	+
.854	260.84	F	55	+++	+	-	+
.853	262.96	M	45	+++	-	++	-
.851	265.76	F	65	+++	+	-	+
.845	266.53	M	45	+++	-	-	+
.842	266.78	F	65	++	+	++	+
.839	267.06	F	55	++	+	++	+
.832	267.40	M	45	++	-	++	+
.827	267.49	M	65	+	-	++	+
.826	267.60	M	55	+	-	++	+
.822	268.19	F	55	++	-	-	+
.820	268.90	F	65	+++	-	-	+
.811	282.99	M	55	+++	+	+	-
.809	283.20	F	65	++	-	++	+
.809	287.43	M	65	+++	+	+	-
.805	287.66	F	55	++	-	++	+
.794	288.78	M	65	++	+	++	-
.790	290.85	M	55	++	+	++	-
.788	291.95	M	65	++	+	+	+
.786	292.06	F	45	+++	+	N	+
.783	293.31	M	55	++	+	+	+
.774	295.13	M	55	+++	-	+	-
.771	295.67	M	65	+++	-	+	-
.768	301.79	F	55	+++	+	++	-
.765	308.51	F	65	+++	+	++	-
.754	309.00	M	65	++	-	++	-
.749	309.37	M	45	+	+	++	+
.749	310.27	M	55	++	-	++	-
.747	310.74	M	65	++	-	+	+
.745	310.80	F	45	+++	-	N	+
.741	311.30	M	55	++	-	+	+
.741	311.86	F	45	+++	+	-	+

Sex M/F = male/female; age 45/55/65 = <50 / 50-59/ ≥60; symptoms = +++ typical ap/ ++ atypical ap / + n.c.p./asymptomatic; T +/- = +/- thallium scintigram; S -/++/+/N = negative stress ECG/ + positive stress ECG with 1-2 mm ST segment depression/++ positive stress ECG with ≥2 mm ST segment depression/ N non-diagnostic test result; +/- C = +/- cine fluoroscopy.

continuation (2) Table 5.1

Post-test prob. %	Cumulative number of patients	Sex	Age	Sympt	T	S	C
.724	313.44	F	55	+++	-	++	-
.724	313.56	F	45	++	+	++	+
.721	315.29	F	65	+++	-	++	-
.717	315.45	F	65	+	+	++	+
.716	315.53	F	55	+	+	++	+
.703	315.65	M	45	+	-	++	+
.694	315.79	F	45	+++	-	-	+
.680	325.16	M	45	+++	+	+	-
.675	325.27	F	45	++	-	++	+
.672	348.18	M	55	+++	+	N	-
.670	358.34	M	65	+++	+	N	-
.668	358.44	F	65	+	-	++	+
.667	358.49	F	55	+	-	++	+
.654	361.15	M	45	++	+	++	-
.645	362.10	M	45	++	+	+	+
.639	365.23	M	65	++	+	N	+
.638	365.49	M	65	+	+	+	+
.635	365.98	M	55	+	+	+	+
.632	368.21	M	55	++	+	N	+
.628	369.46	M	45	+++	-	+	-
.620	372.52	M	55	+++	-	N	-
.618	374.74	F	45	+++	+	++	-
.617	376.09	M	65	+++	-	N	-
.614	396.52	M	55	+++	+	-	-
.611	400.26	M	65	+++	+	-	-
.609	400.43	F	65	++	+	+	+
.603	400.73	F	55	++	+	+	+
.601	401.92	M	45	++	-	++	-
.591	402.35	M	45	++	-	+	+
.584	403.75	M	65	++	-	N	+
.583	403.84	M	65	+	-	+	+
.580	404.00	M	55	+	-	+	+
.579	405.33	M	65	++	+	-	+
.577	406.33	M	55	++	-	N	+
.572	408.64	M	55	++	+	-	+
.563	409.23	F	45	+++	-	++	-
.558	411.99	M	55	+++	-	-	-
.556	412.03	F	45	+	+	++	+
.555	412.54	M	65	+++	-	-	-
.553	412.69	F	65	++	-	+	+
.547	412.97	F	55	++	-	+	+
.530	414.12	M	65	++	+	+	-

Sex M/F = male/female; age 45/55/65 = <50 / 50-59/ ≥60; symptoms = +++ typical ap/ ++ atypical ap/ + n.c.p./asymptomatic; T +/- = +/- thallium scintigram; S -/+ /++/N = negative stress ECG/ + positive stress ECG with 1-2 mm ST segment depression/++ positive stress ECG with ≥2 mm ST segment depression/ N non-diagnostic test result; +/- C = +/- cine fluoroscopy.

Post-test prob. %	Cumulative number of patients	Sex	Age	Sympt	T	S	C
.523	417.55	M	55	++	+	+	-
.522	418.15	M	65	++	-	-	+
.515	419.20	M	55	++	-	-	+
.503	424.34	M	45	+++	+	N	-
.499	424.37	F	45	+	-	++	+
.491	428.00	F	55	+++	+	+	-
.487	430.55	F	65	+++	+	+	-
.472	431.08	M	65	++	-	+	-
.465	431.60	M	45	+	+	+	+
.465	433.18	M	55	++	-	+	-
.465	433.49	F	65	++	+	++	-
.465	434.12	M	45	++	+	N	+
.459	435.09	F	55	++	+	++	-
.456	435.64	M	65	+	+	N	+
.453	436.36	M	55	+	+	N	+
.446	437.08	M	45	+++	-	N	-
.440	454.61	M	45	+++	+	-	-
.437	454.99	M	65	+	+	++	-
.434	455.99	F	55	+++	-	+	-
.434	457.27	M	55	+	+	++	-
.433	457.40	F	45	++	+	+	+
.430	458.10	F	65	+++	-	+	-
.427	458.86	F	65	++	+	N	+
.425	458.99	F	65	+	+	+	+
.423	459.08	F	55	+	+	+	+
.420	459.99	F	55	++	+	N	+
.409	460.17	M	45	+	-	+	-
.409	460.46	F	65	++	-	++	+
.408	460.75	M	45	++	-	N	-
.402	461.66	F	55	++	-	++	+
.402	464.11	M	45	++	+	-	+
.400	464.31	M	65	+	-	N	+
.397	464.56	M	55	+	-	N	+
.394	464.86	M	65	+	+	-	+
.392	465.83	M	55	+	+	-	+
.385	468.30	M	45	+++	-	-	-
.381	468.43	M	65	+	-	++	-
.378	468.87	M	55	+	-	++	-
.378	469.00	F	45	++	-	+	+
.371	469.71	F	65	++	-	N	+
.370	469.79	F	65	+	-	+	+
.368	469.86	F	55	+	-	+	+

Sex M/F = male/female; age 45/55/65 = <50 / 50-59/ ≥60; symptoms = +++ typical ap/ ++ atypical ap / + n.c.p./asymptomatic; T +/- = +/- thallium scintigram; S -/+ ++/N = negative stress ECG/ + positive stress ECG with 1-2 mm ST segment depression/++ positive stress ECG with ≥2 mm ST segment depression/ N non-diagnostic test result; +/- C = +/- cine fluoroscopy.

continuation (4) Table 5.1

Post-test prob. %	Cumulative number of patients	Sex	Age	Sympt	T	S	C
.366	470.15	F	65	++	+	-	+
.366	471.00	F	55	++	-	N	+
.360	471.84	F	55	++	+	-	+
.356	476.25	M	45	++	+	+	-
.350	480.31	M	65	++	+	N	-
.348	481.47	M	45	++	-	-	+
.343	489.63	M	55	++	+	N	-
.341	489.74	M	65	+	-	-	+
.338	490.08	M	55	+	-	-	+
.321	491.42	F	45	+++	+	+	-
.315	501.61	F	55	+++	+	N	-
.314	501.89	F	65	++	-	-	+
.312	512.46	F	65	+++	+	N	-
.309	513.26	F	55	++	-	-	+
.305	515.37	M	45	++	-	+	-
.299	517.32	M	65	++	-	N	-
.299	518.05	F	45	++	+	++	-
.294	519.90	M	65	++	+	-	-
.293	520.16	M	45	+	+	N	+
.293	524.08	M	55	++	-	N	-
.288	533.13	M	55	++	+	-	-
.277	535.57	M	45	+	+	++	-
.273	535.95	F	45	+++	-	+	-
.268	538.87	F	55	+++	-	N	-
.268	538.92	F	45	+	+	+	+
.267	539.05	F	45	++	+	N	+
.264	542.08	F	65	+++	-	N	-
.263	550.58	F	55	+++	+	-	-
.261	551.01	F	65	+	+	N	+
.260	554.64	F	65	+++	+	-	-
.259	554.85	F	55	+	+	N	+
.253	555.56	F	45	++	-	++	-
.249	556.46	M	65	++	-	-	-
.248	556.55	M	45	+	-	N	+
.244	557.89	M	45	+	+	-	+
.243	562.28	M	55	++	-	-	-
.233	563.16	M	45	+	-	++	-
.225	563.20	F	45	+	-	+	+
.225	563.33	F	45	++	-	N	+
.221	565.79	F	55	+++	-	-	-
.220	566.27	F	45	++	+	-	+
.219	566.58	F	65	+	-	N	+

Sex M/F = male/female; age 45/55/65 = <50 / 50-59/ ≥60; symptoms = +++ typical ap/ ++ atypical ap / + n.c.p./asymptomatic; T +/- = +/- thallium scintigram; S -/+ ++/N = negative stress ECG/ + positive stress ECG with 1-2 mm ST segment depression/++ positive stress ECG with ≥2 mm ST segment depression/ N non-diagnostic test result; +/- C = +/- cine fluoroscopy.

continuation (5) Table 5.1

Post-test prob. %	Cumulative number of patients	Sex	Age	Sympt	T	S	C
.218	567.63	F	65	+++	-	-	-
.218	567.79	F	55	+	-	N	+
.215	568.00	F	65	+	+	-	+
.214	568.26	F	55	+	+	-	+
.209	571.80	M	45	++	+	N	-
.204	572.29	M	45	+	-	-	+
.202	572.65	F	65	++	+	+	-
.198	574.43	F	55	++	+	+	-
.184	575.03	M	65	+	+	+	-
.184	576.29	F	45	+++	+	N	-
.183	576.76	F	45	++	-	-	+
.183	579.97	M	55	+	+	+	-
.178	580.13	F	65	+	-	-	+
.177	580.31	F	55	+	-	-	+
.173	582.06	M	45	++	-	N	-
.170	596.94	M	45	++	+	-	-
.168	597.29	F	65	++	-	+	-
.164	599.04	F	55	++	-	+	-
.152	599.26	M	65	+	-	+	-
.152	599.64	F	45	+++	-	N	-
.151	600.83	M	55	+	-	+	-
.149	604.82	F	45	+++	+	-	-
.148	604.86	F	45	+	+	N	+
.148	605.36	F	65	+	+	++	-
.148	606.00	F	55	+	+	++	-
.140	613.42	M	45	++	-	-	-
.122	614.60	F	45	+++	-	-	-
.122	614.63	F	45	+	-	N	+
.122	615.00	F	65	++	-	++	-
.121	615.47	F	55	+	-	++	-
.119	615.66	F	45	+	+	-	+
.111	616.95	F	45	++	+	+	-
.108	619.03	F	65	++	+	N	-
.106	626.04	F	55	++	+	N	-
.101	631.89	M	45	+	+	+	-
.097	633.61	M	65	+	+	N	-
.097	633.75	F	45	+	-	-	+
.096	640.07	M	55	+	+	N	-
.090	641.37	F	45	++	-	+	-
.088	643.44	F	65	++	-	N	-
.086	644.30	F	65	++	+	-	-
.086	651.30	F	55	++	-	n	-

Sex M/F = male/female; age 45/55/65 = <50 / 50-59/ ≥60; symptoms = +++ typical ap/ ++ atypical ap / + n.c.p./asymptomatic; T +/- = +/- thallium scintigram; S -/+ ++/N = negative stress ECG/ + positive stress ECG with 1-2 mm ST segment depression/++ positive stress ECG with ≥2 mm ST segment depression/ N non-diagnostic test result; +/- C = +/- cine fluoroscopy.

continuation (6) Table 5.1

Post-test prob. %	Cumulative number of patients	Sex	Age	Sympt	T	S	C
.084	658.33	F	55	++	+	-	-
.082	660.53	M	45	+	-	+	-
.079	661.21	F	45	+	+	++	-
.079	661.86	M	65	+	-	N	-
.078	664.24	M	55	+	-	N	-
.077	665.28	M	65	+	+	-	-
.076	674.55	M	55	+	+	-	-
.069	675.41	F	65	++	-	-	-
.068	682.46	F	55	++	-	-	-
.064	682.96	F	45	+	-	++	-
.062	683.36	M	65	+	-	-	-
.062	686.86	M	55	+	-	-	-
.056	688.47	F	45	++	+	N	-
.051	692.11	M	45	+	+	N	-
.048	692.84	F	65	+	+	+	-
.048	694.36	F	55	+	+	+	-
.045	695.98	F	45	++	-	N	-
.044	701.96	F	45	++	+	-	-
.041	703.35	M	45	+	-	N	-
.040	723.07	M	45	+	+	-	-
.039	723.62	F	65	+	-	+	-
.039	724.76	F	55	+	-	+	-
.035	730.81	F	45	++	-	-	-
.032	738.33	M	45	+	-	-	-
.025	739.71	F	45	+	+	+	-
.024	742.79	F	65	+	+	N	-
.024	747.08	F	55	+	+	N	-
.020	748.82	F	45	+	-	+	-
.019	750.46	F	65	+	-	N	-
.019	753.70	F	55	+	-	N	-
.018	755.32	F	65	+	+	-	-
.018	760.78	F	55	+	+	-	-
.015	762.01	F	65	+	-	-	-
.015	766.14	F	55	+	-	-	-
.012	767.33	F	45	+	+	N	-
.009	768.23	F	45	+	-	N	-
.009	773.79	F	45	+	+	-	-
.007	778.00	F	45	+	-	-	-

Sex M/F = male/female; age 45/55/65 = <50 / 50-59/ ≥60; symptoms = +++ typical ap/ ++ atypical ap / + n.c.p./asymptomatic; T +/- = +/- thallium scintigram; S -/+ ++/N = negative stress ECG/ + positive stress ECG with 1-2 mm ST segment depression/++ positive stress ECG with ≥2 mm ST segment depression/ N non-diagnostic test result; +/- C = +/- cine fluoroscopy.

Appendix, Chapter 5

Log-linear regression model

A special class of statistical techniques, called log-linear models, has been formulated for the analysis of categorical data. Log-linear models are similar to regression models. These models are useful for uncovering complex relationships between variables in a multiway cross tabulation. All variables, i.e. the diagnostic variables as well as the disease variable, are considered independent variables, while the dependent variable is the number of cases in each cell. Thus, the log-linear model estimates the real number of persons with and without disease for all combinations of the diagnostic characteristics. This in contrast to the logistic regression model as described in Chapter 4, which only gives an estimate of the ratio of the number of diseased and non-diseased persons.

Like the logistic regression model, the relations among the variables for the log-linear model are also assumed to be multiplicative. To obtain a model which is linear, the natural logs of the cell frequencies, rather than the actual counts, have to be used*.

The log-linear model is purely descriptive. The number of persons in a cell is a function of the diagnostic variables as well as the disease variable. To use this function for a probability estimate, the parameters of this function have to be calculated. The coefficients describing the main effects and the interactions between different variables as well as the constant term are estimated from the data by maximizing a likelihood function (see SPSS manual).

When the model contains all main effects and interaction terms, the cell frequencies predicted from the model are exactly the same as observed. Such a model is called a saturated model but is not a parsimonious description of the relationship between the diagnostic variables and the disease probability. It may, however, serve as starting point for exploring an unsaturated model that could be used to represent the data. Many models with different terms are possible for a set of diagnostic variables. As there is a great similarity between the log-linear and logistic regression model**, the same guidelines for model selection in regression analysis also apply to log-linear models.

A model should fit the data, be substantively interpretable, and be as simple (parsimonious) as possible. For example, if models with and without higher-order interaction terms fit the data well, the simpler models are usually preferable since higher-order interaction terms are difficult to interpret.

* Recall that the natural log of a number is the power to which the number e (2.718) is raised to give that number.

** The log-linear model can also be used to estimate the disease odds for all categories by dividing the number of diseased by the number of non-diseased patients for each of the combinations of the diagnostic characteristics. Thus, these estimations of the disease odds can be compared with the results of a logistic regression. However, the logistic regression model only estimates the Odds and uses the exact number of patients for all different diagnostic categories. Therefore, the results will be exactly the same if in the log-linear model the predicted number of patients of all these diagnostic categories are equal to the observed number. This is the case if all higher-order interaction terms between all diagnostic variables, thus with exclusion of the disease variable, are included.

A substantial proportion of the patients was not submitted to thallium scintigraphy and a stress ECG (Table 5A.1). Exempting these patients from the analysis would also possibly lead to selection bias, but would certainly result in a considerable reduction of the study population with loss of the information concerning the other diagnostic variables. For all these patients the result of the concerning test variable was coded as an additional outcome level.

The numbers of diseased and non-diseased patients are also estimated for these additional levels. However, for the construction of the protocol it is necessary to predict for both diseased and non-diseased patients what the test result would have been if the test had been performed.

One way to do this would be to use the distribution of the different test results as observed in the diseased and non-diseased patients in whom these tests have been performed (Table 5A.2). For example for the 122 diseased patients who were not subjected to a stress ECG the result is predicted to be negative in 23.2 patients (=19.2% of 122).

However, as can be seen from Table 5A.1, the patients who did or did not have the test are not comparable. For the patients who did not have a thallium scintigram or a stress ECG the disease probability is higher. This suggests a selection process.

The differences in probability are, at least partly, caused by differences in symptomatology. Thus, patients who did not have a test can only be compared with patients who were submitted to the test concerned and had the same symptomatology.

For each of the combination of clinical characteristics for diseased and non-diseased patients the distribution of the results of the stress ECG as well as the thallium scintigraphy should be analysed. Because of the low number of patients for most of the categories this distribution of the results has also to be estimated by the log-linear model.

To make the predictions more precise, the results of the cine fluoroscopy which was performed in all patients have been used for the estimations of these distributions. The following example will demonstrate this procedure.

In 6 non-diseased females, 50-59 years old with atypical angina and a negative result for the cine-fluoroscopy, a stress ECG was not performed. The log-linear model estimates for this category of patients a distribution of the results of the stress ECG as demonstrated in the first column in Table 5A.3. According to these percentages, these 6 females have been redistributed over the different test results. The outcome is shown in the second column of Table 5A.3.

In conclusion, with help of the log-linear model the number of patients with or without disease is estimated for all combinations of diagnostic characteristics and consecutively the patients in whom not all tests have been performed are redistributed in the way described above. Thus, the log-linear model enables the composition of a list for all combinations of diagnostic variables, ranked by the estimated post-test probability and quantity of each combination. Table 5.1 demonstrates the resulting list.

Table 5A.1 Distributions of diseased (CAD+) and non-diseased (CAD-) patients with or without a thallium scintigram or a stress ECG

Thallium	CAD+	CAD-	Total
Performed	59 (42%)	82 (58%)	141 (100%)
Not performed	364 (57%)	273 (43%)	637 (100%)

Stress ECG	CAD+	CAD-	Total
performed	301 (52%)	279 (48%)	580 (100%)
not performed	122 (62%)	76 (38%)	198 (100%)

Table 5A.2 Distribution of the results of the stress ECG for diseased (CAD+) and non-diseased patients (CAD-)

ST depression	CAD+	CAD-
0 - < 1mm	19.2%	48.7%
1 - < 2mm	17.6%	13.6%
≥ 2mm	38.9%	6.5%
non-diagn.	24.3%	31.2%
Total	100.0%	100.0%

Table 5A.3 Distribution of the stress ECG results for 6 non-diseased females, 50-59 years with atypical angina and a negative fluoroscopy test, as estimated by the log-linear model

ST depression	Percentage	Number
0 - < 1mm	3.6%	0.2
1 - < 2mm	9.8%	0.6
≥ 2 mm	42.7%	2.6
non-diagn.	43.9%	2.6
Total	100.0%	6.0

6

Arteriographic findings, their distribution over gender, age and symptoms, and the diagnostic value of these clinical variables in the diagnosis of coronary artery disease

6.1 Literature review of the value of the clinical variables sex, age and symptomatology

Clinical data that come from history, such as sex, age and symptomatology may also be regarded as "tests", as these also influence the probability of having disease. These "tests" should be used first because they do not cost any money, are hardly a burden to the patient, and as Sackett (1985)^{S1} stated, "a clinical examination is far more powerful than a laboratory evaluation in establishing diagnoses".

This is also the opinion of clinicians such as Ascoop (1979)^{A4} who emphasised the importance of the diagnostic value of history taking by an experienced clinician as compared with the value of the stress ECG. Bruschke (1970)^{B3} concluded the same in his thesis, comparing history, ECG and stress ECG with arteriogram findings.

But, stressing the importance of history taking, Bruschke also warned that the evaluation of the symptoms may be difficult, even with the use of questionnaires filled out by the patient. Proudfit (1966)^{P7} noted that the terminology used in history taking may also cause problems; for example, atypical angina may have a different meaning for each clinician. Furthermore, the classification of symptoms may be difficult due to a patient's emotional colouring of the illness. Many patients were referred to his clinic with angina at rest and at exercise, often patients with difficult therapeutic problems. Patients with this type of symptomatology had been hospitalised on numerous occasions. Thus the finding of normal coronary arteries was common in this subgroup.

Attempts to obtain answers to the various questions can be complicated by problems inherent in the physician-patient relationship: namely, the failure of the physician to ask the proper questions and to pursue the answers, and the failure of the patient to reply adequately, whether due to a disordered mental state or merely due to an inability to remember accurately or to verbalise properly. However, the importance of taking an adequate history cannot be over-emphasised (Goodwin, 1991)^{G3}. In this chapter, first the informative value of the different clinical variables as provided by different sources from the literature, will be discussed. This will enable a comparison with the informative value of non-invasive tests.

Furthermore, the predictive value of non-invasive tests is dependent on the estimation of the pre-test probability. This estimation of the pre-test probability will be more precise if the freely available information of the clinical variables is used. As discussed in Chapter 4, the informative value of tests may be analysed using the LR. Calculating this test parameter for sex, age and symptomatology enables a comparison of the diagnostic value of these clinical variables as well as a comparison with the diagnostic value of non-invasive tests. To calculate the LRs of the clinical variables, the results of two published studies by Diamond in 1979 and Chaitman in 1981 were used (Diamond, 1979^{D5}; Chaitman, 1981^{C2}). Diamond collected the symptomatology and the findings of the coronary arteriograms of 4,952 symptomatic patients, from published studies in the literature. Although he made a distinction between the symptoms, typical angina, atypical angina, and non-anginal chest pain, the criteria for these categories were not further defined in his study. Significant coronary artery disease was defined as having a lesion of at least a 75% loss in cross-section area in a major vessel. It is also unclear whether or not patients with a myocardial infarction were included. Furthermore, he obtained data on age and sex and the prevalence of coronary artery disease as detected at autopsy in 23,996 asymptomatic subjects. Using this data, the LRs can be calculated for both sexes and for the different levels of age and symptomatology (Table 6.1).

Table 6.1 LR values for clinical variables, adapted from Diamond and Chaitman (Diamond, 1979; Chaitman, 1981)

	Diamond	Chaitman
Sex		
male	1.23	1.76
female	0.45	0.34
Age		
30 - 39	0.20	0.30
40 - 49	0.61	0.65
50 - 59	1.23	1.16
60 - 69	1.77	2.07
Symptoms		
typical angina	5.28	8.61
atypical angina	0.65	1.32
non-anginal chest pain	0.12	0.15

Chaitman obtained data from the Coronary Artery Surgery Study (the CASS study) and provided estimations of the post-test probability of having coronary artery disease based on the clinical variables, symptoms, age and sex (Table 6.2). He compared coronary arteriographic results with the presence of symptoms (definite angina, probable angina and nonspecific chest pain) in 8,157 patients, also differing in age and sex. Patients with "unstable angina" and myocardial infarction were excluded. Although significant coronary artery disease was defined differently in Chaitman's study, the data provided allowed adaptation to the criteria used in the present study.

The LR_s for the clinical variables can also be calculated from these data and the values can be compared with those calculated from Diamond's data as demonstrated in Table 6.1. The values for the LR appear to be comparable only for the age categories. The LR values differ considerably for both sexes and certainly for the symptom categories and these discrepancies will be discussed at the end of this chapter.

Apart from the evaluation of the diagnostic value with help of the LR, the objective of the analysis should also be to obtain an estimation of the probability of coronary artery disease given the sex, a certain age and symptom, as this probability is of importance for the clinical management of patients.

As shown in Chapter 4, this probability analysis can be conducted with help of the LR. However, there is an additional problem here, as the three "tests" are performed in parallel. When the diagnostic chain is short, i.e. consists of no more than three tests, Sackett suggests using LR_s for the estimation of the posterior probability assuming independence between the tests. The formula for the Odds of coronary artery disease for a patient with a certain sex, age and symptom is then as follows:

$$\text{posterior Odds} = \text{LR}(\text{sex}) \times \text{LR}(\text{age}) \times \text{LR}(\text{symptom}) \times \text{prior Odds}$$

where Odds = probability of CAD/(1-probability of CAD)

With help of this formula and using the LR_s from Diamond's study as well as from Chaitman's, the posterior probabilities can be estimated (Table 6.2).

It is, however, questionable to use this formula because this presumes statistical independence between symptoms, age and sex. Neither this possible mutual dependence nor its implication were discussed in Diamond's publication. In contrast with Diamond's data, Chaitman's data on sex, age and symptoms come from one population. As the percentage of patients with coronary artery disease is presented for each combination of clinical characteristics, it is possible to estimate the probability of disease for all the combinations of clinical variables directly and with help of the LR_s (Table 6.2). As shown before (Table 6.1) the LR_s calculated from Diamond's data and from Chaitman's data differ considerably, so it can be foreseen that the estimations based on these LR_s will also differ. These predicted differences are indeed demonstrated in Table 6.2. The discrepancy between the LR_s and consequently the estimations for the posterior probabilities can be due to differences in the definitions used for the symptomatology and dissimilarity in the composition of the populations; Diamond's data come partly from autopsies of asymptomatic persons and furthermore, it is unclear whether or not patients with a previous myocardial infarction were included. From Chaitman's data the posterior probabilities can be estimated directly and by calculation using LR_s. The results of both methods do not diverge much (Table 6.2); or perhaps only slightly for a number of female categories. So Sackett may be right when he stated that for short diagnostic chains independence may be assumed and LR_s may be used for the estimation of the posterior probability. Granted that these differences are minimal, it still seems advisable, certainly when more diagnostic variables are involved, to investigate whether the assumption of independence is appropriate.

Table 6.2 Post-test probability estimates in % calculated with help of LR_s as adapted from Diamond (1), calculated with help of LR_s as adapted from Chaitman (2), and from direct estimates, adapted from Chaitman (3) (Diamond, 1979; Chaitman, 1981)

	Diamond ¹	Chaitman ²	Chaitman ³
Male			
Typical angina			
30 - 39	67	85	86
40 - 49	86	92	89
50 - 59	92	96	97
60 - 69	95	98	96
Atypical angina			
30 - 39	20	47	50
40 - 49	43	65	62
50 - 59	60	77	77
60 - 69	69	86	83
Nonspecific chest pain			
30 - 39	5	9	7
40 - 49	13	18	19
50 - 59	23	28	30
60 - 69	30	41	42
Female			
Typical angina			
30 - 39	42	53	70
40 - 49	70	71	60
50 - 59	82	81	76
60 - 69	87	89	84
Atypical angina			
30 - 39	8	15	28
40 - 49	22	27	35
50 - 59	36	40	36
60 - 69	45	54	55
Nonspecific chest pain			
30 - 39	2	2	4
40 - 49	5	4	6
50 - 59	10	7	10
60 - 69	13	12	21

6.2 Arteriographic findings in relation to sex, age and symptoms

The severity of disease can be described on several ways. One way is to express the extent of severity in one, two and triple vessel disease. Another way of expressing the severity of lesions of the coronary arteries is one based on 5 categories of arteriographic findings. Such a division for our study population is demonstrated in Table 6.3. In the following results significant coronary artery disease is defined as minimally one lesion as described in the categories D and E. Thus, 423 patients met these criteria and the prevalence of disease in our study population was 54.4%.

The prevalence of coronary artery disease differs markedly per stratum of the clinical variables (Table 6.4).

Table 6.3 Subdivision of arteriographic findings in 5 categories for the study population (N=778)

Category 1*	23%
Category 2*	6%
Category 3*	16%
Category 4*	8%
Category 5*	46%
Total	100% (N=778)

* Category 1: No coronary artery disease at all; Category 2: No coronary artery disease but only a muscle bridge; Category 3: Only minimal coronary artery disease, no significant stenosis present; Category 4: Coronary artery disease, with at least one lesion of 75% loss in cross-sectional area (moderate stenosis) and supplying 1.5 myocardial units, but with no lesions of 90%; Category 5: Coronary artery disease, with at least one lesion of $\geq 90\%$ loss in cross-sectional area (severe stenosis), and the artery supplying 1.5 myocardial units.

Table 6.4 Prevalence of significant coronary artery disease (CAD+) and the total number of patients (n) studied, results stratified for different clinical variables

	CAD+ (%)	n
Female	36.8	250
Male	62.6	528
Asymptomatic	14.5	55
Nonspecific chest pain	6.3	80
Atypical angina	29.6	179
Typical angina	76.0	464
< 50 years	37.5	224
50 - 59 years	55.4	339
≥ 60 years	70.2	215

The probability of disease is, as expected, higher in males than females, and increases with age and with a more typical anginal pattern. Using the cumulative distribution of the myocardial score as described in Chapter 3, it is possible to obtain information about the severity and/or the extent of coronary artery disease for these categories. The cumulative distribution distribution in Figure 6.1 shows that the severity increases with a more typical anginal pattern as well as with increasing age, and is more extensive in males as compared with females. Thus, severity and prevalence demonstrate the same pattern over the strata.

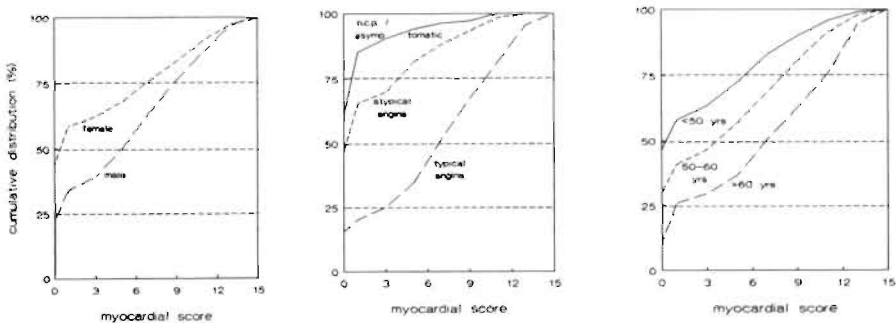


Figure 6.1 Cumulative distribution of the myocardial score for gender, symptomatology, and age groups. The vertical axis represents the percentage of patients with a myocardial score equal to or less than the corresponding value on the horizontal axis, e.g. 25% of the patients with typical angina have a myocardial score of 3 or less.

The distribution of the patients with and without disease and stratified for sex, age and symptomatology is shown in Table 6.5. More than half of the patients investigated were males in the age group 40-59. In this table four categories of symptomatology are shown; in the following analysis the categories nonspecific chest pain and asymptomatic have been combined in one category n.c.p./asymptomatic. As discussed in section 6.1 clinical data that come from history, such as sex, age and symptomatology may also be regarded as "tests", as these also influence the probability of having disease.

Table 6.5 Composition of study population, consisting of 778 patients, categorised by sex, age, and symptomatology and arteriogram findings

Sex	Age	Symptom	CAD		Total
			CAD+	CAD-	
Male	< 50 yr	typical angina	64	17	81
		atypical angina	11	34	45
		nonspecific chest pain	2	29	31
		asymptomatic	2	14	16
		Total	79	94	173
	50-59 yr	typical angina	126	28	154
		atypical angina	19	24	43
		nonspecific chest pain	2	13	15
		asymptomatic	3	13	16
		Total	150	78	228
	≥ 60 yr	typical angina	91	7	98
		atypical angina	10	12	22
		nonspecific chest pain	0	0	0
		asymptomatic	1	6	7
		Total	102	25	127
Female	< 50 yr	typical angina	3	11	14
		atypical angina	2	19	21
		nonspecific chest pain	0	10	10
		asymptomatic	0	6	6
		Total	5	46	51
	50-59 yr	typical angina	28	23	51
		atypical angina	7	31	38
		nonspecific chest pain	1	16	17
		asymptomatic	2	3	5
		Total	38	73	111
	≥ 60 yr	typical angina	45	21	66
		atypical angina	4	6	10
		nonspecific chest pain	0	7	7
		asymptomatic	0	5	5
		Total	49	39	88

CAD +/- = +/-significant coronary artery disease; Symptom as defined in Chapter 2

The informative value of these tests may be analysed using the LR as a test parameter; this was elaborated in Chapter 4. In Table 6.6 this test parameter is calculated with its confidence intervals for the different categories of gender, age

and symptomatology. Although the LRs do not differ much from 1, indicating that the discriminating value is not high, this information should still be used to get a more precise probability estimate as it is freely available. Furthermore, these values of the LRs can be compared with the results of two published studies by Diamond in 1979 and Chaitman in 1981 as presented in the previous section. Except for the value of typical angina which is rather low in our population, the values of the LR of the other categories correspond well with the findings of Chaitman and Diamond.

Table 6.6 LR values for the different categories of the clinical variables

	CAD+	CAD-	LR	95%-C.I.
Sex				
male	331	197	1.41	1.18-1.68
female	92	158	0.49	0.38-0.63
Age				
< 50 yr	84	140	0.50	0.38-0.66
50 - 59 yr	188	151	1.04	0.84-1.29
≥ 60 yr	151	64	1.98	1.48-2.65
Symptoms				
typical angina	357	107	2.80	2.26-3.48
atypical angina	53	126	0.35	0.26-0.49
n.c.p./asymptomatic	13	122	0.09	0.05-0.16

95%-C.I. = 95%-Confidence Interval of the LR

Electrocardiographic stress testing and its value in the detection of coronary artery disease

7.1 Literature review

7.1.1 Introduction

A comprehensive account of exercise testing is beyond the scope of this study: only its value in diagnosing coronary artery disease, especially in patients with symptoms, and its role in patient selection for diagnostic coronary angiography will be discussed. (references used: Cohn P, 1985^{C3}; Ascoop, 1979^{A5} and Goldschlager, 1982^{G4}). The exercise ECG test is a non-invasive and relatively inexpensive test which can also be performed outside a hospital environment. It is a safe test if performed under certain conditions. Stress testing can be performed in different ways as explained and compared by Fortuin (1977^{F3}). The bicycle and treadmill are the commonest devices used in the performance of stress testing. Certain differences should be pointed out. The treadmill may yield more reproducible data, as the subject has no voluntary control over the workload. With regard to the bicycle incoordination and leg fatigue, limiting exercise capacity is less likely to occur with Dutch patients as compared with the American population, as most Dutch are familiar with bicycle pedalling. Surface electrode noise may be less than on a treadmill, allowing better recording and the bicycle ergometer is also less expensive. The physiological differences observed are minimal (Cohn P, 1985^{C3}).

7.1.2 ST segment change

A large number of ECG changes following stress testing have been described in patients with ischaemic heart disease, but only changes in the ST segment have been shown to be of importance (Fortuin, 1977^{F3}; Goldschlager, 1976^{G5} and 1982^{G4}; Ascoop, 1979^{A6} and Cohn P, 1985^{C3}). The most useful changes are J point depression of at least 1 mm with downsloping or horizontal (square wave) ST segments of at least 80 msec; slowly upsloping ST segment depression, defined as 2 mm of ST depression measured 80 msec from the J point (Rijneke, 1980^{R4}; Van Tellingen, 1984^{T3}); and ST segment elevation. The magnitude of ST segment depression is measured by some workers at the J point and 80 msec after the J point by others (Ellestad, 1977^{E2}), which makes some difference if the ST segment is not horizontal. Whereas ST segment depression indicates non-transmural ischaemia, ST segment elevation often connotes more severe degrees of ischaemia reflecting transmural injury. Goldschlager (1976^{G5}) has indicated that the nature of the ST segment slope is predictive for the extent of coronary disease

shown arteriographically, with downsloping ST depression indicating severe two- and three vessel disease more often than with either horizontal or slowly upsloping ST depression. ST segment elevation indicates high-grade, usually proximal, arterial obstruction in patients without previous myocardial infarction. The rate of depression of the ST segment with increasing heart rate during exercise has been claimed to predict the extent of coronary artery disease (Elamin, 1982^{E3}), but this has been refuted by Thwaites (Thwaites, 1986^{T4}). ST segment depression is a functional sign of an abnormal metabolic supply-demand relationship at the myocardial cell membrane. It does not indicate the cause. Other causes apart from coronary artery disease include anaemia, myocardial hypertrophy, hypoxaemia, treatment with digitalis and hypokalaemia. Myocardial ischaemia due to coronary arterial disease can be provoked by stress ECG, if there is an obstruction of haemodynamic significance of at least 75% loss in cross-section area and if there is no adequate collateral flow (Ascoop, 1979^{A5}).

Stress ECG testing is based on the premise that myocardial ischaemia can be measured, but this is a dynamic, physiological function related to inadequate delivery of oxygen to the myocardium. Similarly, the demonstration of morphological vascular pathology by coronary angiography is not directly indicative of the physiological phenomenon of ischaemia; coronary artery narrowing does not indicate that the myocardium is actually ischaemic, but merely implies that the probability exists (Goldschlager, 1976^{G5}).

Rifkin (1977)^{R5} was the first to evaluate the diagnostic value of exercise-induced ST segment depression in the detection of coronary artery disease with the likelihood ratio as parameter and he suggested values for different degrees of ST segment depression based on a literature study. He analysed exercise-induced ST segment depressions as reported in 3 previous studies involving 100 patients (Martin, 1972^{M3}), 350 patients (Bartel, 1974^{B8}) and 410 patients (Goldschlager, 1976^{G5}), where ST segment depressions were compared with coronary arteriogram findings. Patients with abnormal ECG findings at rest were excluded in these populations, but there was no mention of whether patients with a history of a previous myocardial infarction were excluded.

The ST depressions in these ECG data were re-examined quantitatively and not, as previously done, merely as positive if the ST depression was 1 mm or more or negative if it was less. The values of the ST segment depressions were divided into groups in increments of 0.5 mm. Weighted mean values for sensitivity and specificity (and thus the likelihood ratios) were calculated for all these ST segment depression ranges (Table 7.1). No stratification was made for sex or symptomatology.

Rifkin combined the likelihood ratios of each of these ST segment depression ranges with a pre-test probability percentage based on symptoms, providing post-test probabilities. Also, Diamond (1979^{D5}) published data of likelihood ratios for different subsets of ST segment depression, based on a number of studies reported in the literature (Table 7.1). There was no statement on whether patients with a previous myocardial infarction were excluded and there was no stratification for sex or symptomatology. The LR values for the different strata comparing the results of Rifkin and Diamond differ; an explanation cannot be offered, but as each author compiled the data from different literature sources, these differences may not be too surprising.

Table 7.1 Likelihood ratios for different degrees of ST segment depression (from Rifkin, 1977 and from Diamond, 1979)

Degree of ST segment depression	Likelihood ratio test result from Rifkin	Likelihood ratio test result from Diamond
0 mm - < 0.5 mm	0.28	0.22
0.5 mm - < 1 mm	0.41	0.63
1 mm - < 1.5 mm	2.35	2.1
1.5 mm - < 2 mm	9.3	4.1
2 mm - < 2.5 mm	21.3	11.1
2.5 mm and >	no data	39.0

The ranges of sensitivity and specificity values for stress ECG testing are also broad (Table 7.2) (Simoons, 1976^{S5}; Fortuin, 1977^{F3}; Ascoop, 1979^A; and Cohn, 1985^{C3}). There is some controversy over the reasons for the variations of reported values for the sensitivity and specificity of the electrocardiographic response of the ST segment to exercise in determining the presence or absence of coronary artery disease. Differences in defining the symptomatology, the stress ECG result, and the arteriographic findings, as well as differences in technique of performing the stress ECG and arteriogram, and differences in population composition may all have accounted for the variation. Philbrick (1980)^{P4} proposed various causes which may lead to the differences in reported sensitivity and specificity of stress ECG testing, as already discussed in Chapter 4.

7.1.3 Conclusion

Many studies have appeared correlating stress ECG testing with coronary arteriography in the diagnosis of coronary artery disease. There is a wide range in values of the different test parameters. Patients with a previous myocardial infarction are not always excluded from the study populations, a necessity for evaluating any test for diagnosing coronary artery disease.

Table 7.2 Range of sensitivity and specificity values for stress ECG testing

Author (%)	Year	Cases (n)	Sensitivity (%)	Specificity
McNeer ^{M4}	1978	1472	61	95
Chaitman ^{C4}	1979	200	84	72
Weiner ^{W3}	1979	2045	79	69
Aldrich ^{A6}	1979	181	40	92
Raffo ^{R6}	1979	100	91	96
Borer ^{B9}	1979	75	63	95

7.2 Results of stress ECG testing

7.2.1 Likelihood Ratio according to degree of ST segment depression

Stress ECG testing was performed in 580 out of 778 patients. The results described in this chapter are thus based on the findings in these 580 patients. The relationship between the result of the stress ECG and the result of the gold standard, the cine coronary arteriogram, can be described in an easy and direct way by presenting the percentages of the different test results for patients with and

without coronary artery disease (Table 7.3). The degree of ST-segment depression is a measure of the severity of ischaemia of the myocardium during exercise. The probability of the presence, as well as the severity, of coronary artery disease increases with the more severe forms of ST segment depression. A relatively large proportion (14%) of patients with coronary artery disease have a normal stress ECG. Patients with coronary artery disease who have not exercised sufficiently and whose stress ECG is categorised as non diagnostic are also a difficult group to evaluate. It is not clear whether the stress ECG would have been positive if these patients had exercised adequately (the norm is a heart frequency which has to be reached and which is age-dependent). In patients without coronary artery disease, no patho-physiological phenomena of the myocardium are expected during exercise. However, 15% of the patients without coronary artery disease had a positive stress ECG test, the majority demonstrating an ST segment depression of 1-2 mm. Apparently, patho-physiological phenomena do occur in the myocardium during exercise although the major coronary arteries may be patent. Furthermore, 31% of the stress ECGs of patients without coronary artery disease are non-diagnostic.

The likelihood ratio (LR) of a positive or negative test result is the quotient of the percentages of that test result in patients with or without coronary artery disease. The greater the ST segment depression the higher the LR and the better the discriminating value between patients with and patients without coronary artery disease. As already discussed in Chapter 4, the LR as test parameter has the advantage that for instance a LR can still be calculated for non-diagnostic test results. A non-diagnostic test result with an LR of 0.78 was obtained in 160 out of 580 patients (27%) (Table 7.3). This implies that a non-diagnostic test result provides also information about the probability of having coronary artery disease; it lowers the chance, and the test result has an informative value.

Table 7.3 Stress ECG testing in patients with (CAD+) and without significant coronary artery disease (CAD-) and the likelihood ratio (LR) for the test results

Degree ST segment depression	Number of patients		LR	95%-C.I.
	CAD+	CAD-		
Nil	43 (14.3%)	109 (39.1%)	0.37	0.27-0.50
< 1 mm	15 (5.0%)	27 (0.9%)	0.51	0.27-0.90
1 - < 2 mm	53 (17.6%)	38 (13.6%)	1.29	0.87-1.86
≥ 2 mm	117 (38.9%)	18 (0.6%)	6.02	3.81-9.73
Non-diagnostic	73 (24.3%)	87 (31.2%)	0.78	0.60-1.02
	301 (100%)	279 (100%)		

95%-C.I. = 95% confidence interval of the LR

The results of stress ECG testing with stratification for sex, age and symptomatology (with confidence intervals), are shown in Table 7.4. The 95% confidence intervals for the LRs were calculated with the formula for the approximate interval estimation of the risk ratio with cumulative incidence data. In order to reduce the number of categories without data or insufficient data, data have been merged. Patients with nonspecific chest pain have been combined with the asymptomatic patients. Patients have been grouped into 3 age categories: < 50 years; 50-59 years; ≥ 60 years.

Table 7.4 Stress ECG testing in patients with (CAD+) and without significant coronary artery disease (CAD-), and the likelihood ratio (LR); results stratified for sex, symptoms and age

	Number of patients		LR	95%-C.I.
	CAD+	CAD-		
Female				
Nil	5	41	0.25	0.10- 0.59
< 1 mm	4	13	0.64	0.22- 1.78
1 - < 2 mm	6	17	0.73	0.30- 1.76
≥ 2 mm	22	13	3.50	1.90- 6.45
non-diagnostic	23	40	1.19	0.81- 1.84
Male				
Nil	38	68	0.36	0.26- 0.52
< 1 mm	11	14	0.51	0.22- 1.00
1 - < 2 mm	47	21	1.44	0.88- 2.22
≥ 2 mm	95	5	12.22	5.18-29.9
non-diagnostic	50	47	0.68	0.48- 0.96
N.c.p./asymptomatic				
Nil	5	42	1.13	0.57- 2.10
< 1 mm	0	7	?	?
1 - < 2 mm	0	17	?	?
≥ 2 mm	2	6	3.17	0.74-13.8
non-diagnostic	3	23	1.24	0.47- 3.61
Atypical angina				
Nil	11	40	0.66	0.38- 1.19
< 1 mm	1	13	0.18	0.02- 1.37
1 - < 2 mm	6	15	0.96	0.40- 2.30
≥ 2 mm	9	5	4.31	1.53-12.1
non-diagnostic	16	30	1.28	0.76- 2.01
Typical angina				
Nil	27	27	0.33	0.21- 0.54
< 1 mm	14	7	0.65	0.25- 1.31
1 - < 2 mm	47	6	2.56	1.05- 4.74
≥ 2 mm	106	7	4.95	2.44-10.3
non-diagnostic	54	34	0.52	0.37- 0.76
< 50 years				
Nil	17	63	0.50	0.33- 0.79
< 1 mm	1	9	0.21	0.03- 1.57
1 - < 2 mm	11	21	0.97	0.49- 1.86
≥ 2 mm	26	10	4.83	2.56- 9.54
non-diagnostic	8	14	1.06	0.41- 2.02
50 - 59 years				
Nil	19	40	0.39	0.25- 0.65
< 1 mm	10	14	0.59	0.26- 1.21
1 - < 2 mm	27	13	1.72	0.90- 2.96
≥ 2 mm	46	6	6.36	2.86-14.6
non-diagnostic	39	44	0.74	0.53- 1.10
≥ 60 years				
Nil	7	6	0.54	0.17- 1.21
< 1 mm	4	4	0.46	0.12- 1.72
1 - < 2 mm	15	4	1.74	1.80- 5.08
≥ 2 mm	45	2	10.44	2.63-40.8
non-diagnostic	26	29	0.42	0.29- 0.64

95%-C.I. = 95% confidence interval of the LR

The stratified LR for a ST segment depression of < 1 mm is lower in patients < 50 and ≥ 60 years old and in patients with atypical angina than the LR for a negative test result. The number of patients without coronary artery disease is

thus relatively high in these categories and an explanation cannot be offered. The confidence intervals for a negative test result and a ST segment depression of < 1 mm overlap here and thus the discrepancy can be explained by chance. This may be a reason to interpret an ST segment depression of < 1 mm as negative and only a greater ST segment depression as positive. This, in fact, is done in clinical practice as judged from the literature. Furthermore, if one stratifies for the various clinical variables, the LR of a non-diagnostic test result ranks differently in relation to the other ECG results. Its value fluctuates for the different categories from 0.42 to 1.28, thus reducing or increasing the probability of having disease but no explanation for this can be offered. However the use of beta-blocker medication may be of influence here and may confuse the results. The possible effect of beta blocker medication on the stress ECG result will be further discussed in the last paragraph of this chapter.

The result of the stress ECG is expressed as a ratio variable and the discriminating value of this test can therefore also be illustrated by means of a ROC-curve. The non-diagnostic test result cannot be used for this ROC-curve. The more the curve approaches the diagonal, the less informative the test is. This is the case for females, patients with less specific complaints and younger patients (Figure 7.1). The ROC-curve for the total population is almost the same as that for patients with typical angina (Figure 7.1). The ROC-curves for patients with atypical angina and nonspecific complaints lie significantly lower. This implies that the ROC-curve for the total population cannot be regarded as an "average" of the curves corresponding to the three symptom categories. Symptomatology modifies the discriminating value of the result of the stress ECG. The position of the curve for the total population (crude ROC-curve) is eccentric relative to the stratified curves, illustrating a previously discussed mechanism called collinearity. Furthermore, as can be seen from the different positions of the curves for the three different symptom categories, interaction is also present.

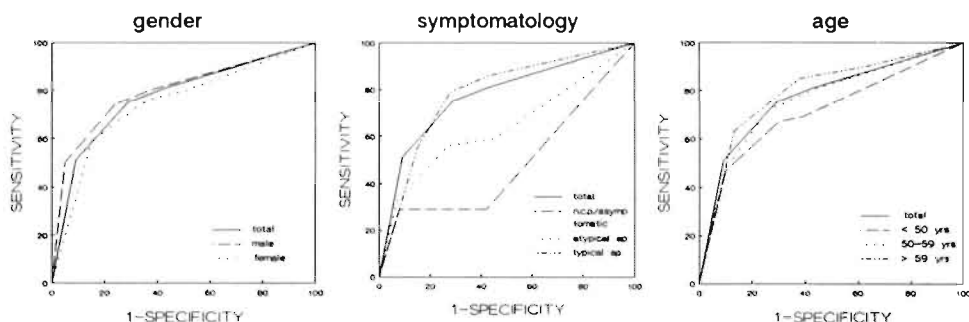


Figure 7.1 ROC curves for stress ECG stratified for gender, symptomatology and age

7.2.2 Influence of clinical variables

As stated before (Chapter 4), to analyse the influence of the clinical variables age, sex and symptoms on the association between the test result and the disease probability and to check for the presence of collinearity and interaction, the (diagnostic) odds ratio (OR) can be used as test parameter, as concluded from Hosmer (Hosmer, 1989^{H4}) and Bouter (Bouter and Van Dongen, 1988^{B10}). Our data were analysed accordingly. Stratification according to sex, age and symptoms demonstrates the presence of collinearity, as the stratified ORs differ from

the crude OR (i.e. calculated for unstratified data), as well as the presence of interaction since the stratified ORs are different per stratum (Table 7.5).

Table 7.5 Crude odds ratio, odds ratio stratified for a single clinical variable and the adjusted odds ratio

Not stratified	OR _{crude}	95%-C.I.	OR _{adjusted}	95%-C.I.
Nil	1	0.68- 2.9	1	
< 1 mm	1.4	0.68- 2.9	0.84	0.35- 2.0
1 - < 2 mm	3.54	2.0 - 6.1	2.66	1.32- 5.34
≥ 2 mm	16.5	8.96-30.3	9.93	4.66-21.2
non-diagnostic	2.13	1.33- 3.41	1.30	0.71- 2.38
Stratified	OR _{single}	95%-C.I.		
Female				
Nil	1			
< 1 mm	2.52	0.59-10.81		
1 - < 2 mm	2.89	0.78-10.78		
≥ 2 mm	13.87	4.38-44.0		
non-diagnostic	4.72	1.63-13.62		
Male				
Nil	1			
< 1 mm	1.4	0.58- 3.4		
1 - < 2 mm	4	2.09- 7.67		
≥ 2 mm	34.0	12.72-90.87		
non-diagnostic	1.9	1.09- 3.34		
N.c.p./asymptomatic				
Nil	1			
< 1 mm	?			
1 - < 2 mm	?			
≥ 2 mm	2.8	0.44-17.8		
non-diagnostic	1.1	0.24- 5.0		
Atypical angina				
Nil	1			
< 1 mm	0.28	0.04- 2.38		
1 - < 2 mm	1.45	0.46- 4.63		
≥ 2 mm	6.55	1.82-23.56		
non-diagnostic	1.94	0.79- 4.78		
Typical angina				
Nil	1			
< 1 mm	2.0	0.7 - 5.73		
1 - < 2 mm	7.83	2.87-21.63		
≥ 2 mm	15.14	5.96-38.48		
non-diagnostic	1.59	0.8 - 3.15		
< 50 years				
Nil	1			
< 1 mm	0.42	0.05- 3.49		
1 - < 2 mm	1.94	0.79- 4.80		
≥ 2 mm	9.64	3.9 -23.82		
non-diagnostic	2.12	0.76- 5.88		
50 - 59 years				
Nil	1			
< 1 mm	1.50	0.57- 4.0		
1 - < 2 mm	4.37	1.86-10.31		
≥ 2 mm	16.1	5.87-44.36		
non-diagnostic	1.87	0.93- 3.74		
≥ 60 years				
Nil	1			
< 1 mm	0.86	0.15- 5.0		
1 - < 2 mm	3.21	0.68-15.16		
≥ 2 mm	19.29	3.23-11.5		
non-diagnostic	0.77	0.23- 2.58		

OR_{single} = stratified for a single variable; OR_{adjusted} = stratified for all variables with logistic regression;
95% C.I. = 95% confidence interval of the OR

An analysis correcting for all possible interactions between the 3 clinical variables is not possible. A number of categories lack data. An adjusted OR has been obtained using a logistic regression model. The result derived, now corrected for the 3 confounders, is also shown in Table 7.5. It appears that there are ostensible differences especially for test results with higher degrees of ST segment depression. As discussed in section 7.1, studies in the literature reporting on stress ECG testing provide no useful information for our study and comparison is not possible because the data provided are not stratified enough. Furthermore, differences in study populations, due to differences in the definition of the symptomatology and in the stress ECG and cine coronary arteriographic results, due to differences in stress ECG and angiography protocols, as well as differences in population composition (e.g. patients with a previous myocardial infarction included or excluded), hinder comparison.

7.2.3 Probability analysis

The presence of collinearity as well as interaction imply that the assumption of independence is not correct. The usually recommended method of Bayes for calculating the post-test probability would involve the use of conditional LRs (or sensitivity and specificity) which is not practical for the number of clinical variables as included in this analysis.

Therefore, a logistic regression model was used here also to estimate the post-test probability for all combinations of sex, age and symptom groups. It is in the probability range 10-20% where the decision has to be made whether or not to perform an additional (more expensive) non-invasive test or even an arteriogram. If only stress ECG testing is used as the non-invasive test, the test result based on independence or dependence could influence the decision for referral for an arteriogram for a number of female categories, for males older than 59 years with atypical angina and with a stress ECG result with a < 1 mm ST segment depression, (Figure 7.2).

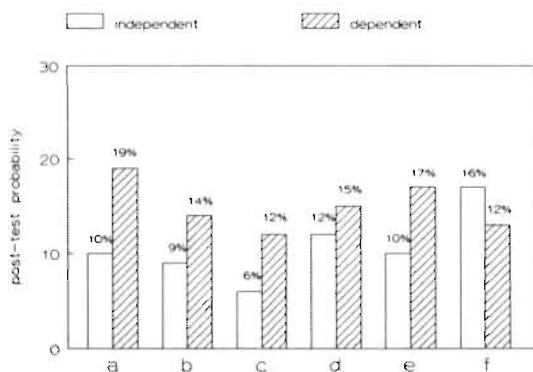


Figure 7.2 Some examples of differences between the post-test probability assuming independence and the post-test probability corrected for dependence, which may lead to different management of patient.

(a) females, < 50 years, n.c.p./asymptomatic, stress ECG nil; (b) females, < 50 years, n.c.p./asymptomatic, stress ECG positive with a ST segment depression of 1-2 mm; (c) females, 50 - 60 years, n.c.p./asymptomatic, stress ECG nil; (d) females, 50 - 60 years, n.c.p./asymptomatic, stress ECG non-diagnostic; (e) females, ≥ 60 years, n.c.p./asymptomatic, stress ECG nil; (f) males, ≥ 60 years, atypical angina, stress ECG positive with a ST segment depression of < 1 mm

Differences between estimation of the post-test probability by logistic regression and estimation without taking dependence into account are present for all categories (addendum 7.1.) As discussed before, these differences are of importance since they may lead to a different management of the patient. As demonstrated in Figure 7.2, there are indeed differences in the 10-20% post-test probability range, leading to a different management of the patient.

In an independent model, the diagnostic variables all contribute independently to the estimate of the probability of coronary artery disease. This means that the estimate of the post-test probability for a patient with all prognostically favourable diagnostic variables will be too low or contrarily too high in the case of all unfavourable characteristics, compared with the estimates using the dependent model. A female of young age with n.c.p./asymptomatic and a negative stress ECG has all these factors reducing the estimate for the post-test probability. As shown in Figure 7.2, the estimation for the probability assuming independence is lower than the estimate accounting for dependence. The differences in the other categories of this figure can all be explained in the same way except for category f. The latter category comprises males of ≥ 60 years with atypical angina and a positive stress ECG and demonstrates a higher value for the post-test probability when assuming independence, as this category has mainly unfavourable factors.

7.2.4 Possible Bias by selective referral

As has been described in Chapter 2, the results can be biased by selective referral for coronary arteriography. First, there are patients who have not been referred for arteriography following stress ECG testing, for instance patients with a nil stress ECG. There are also patients who had an arteriogram without a stress ECG test. The following serves to demonstrate the degree of this type of selection bias. The disease prevalence in patients who had a stress ECG was 52%, and for patients who had been referred for arteriography without a stress ECG it was 62%. This difference in prevalence and possible difference in severity of disease can, however, be based on differences in sex, age or symptomatology between both populations. These clinical variables are, at least partly, the criteria for referral for arteriography without a preceding stress ECG. It is possible to correct for these differences. The prevalence of coronary artery disease in the group without stress ECG testing was 56%, when correction was made for sex, age and symptoms and the difference was then reduced. Complete correction is only possible if all factors which influence the referral pattern of the cardiologist are known and can be taken into account in the analysis. By using stratification in the analysis of the data, one can correct for the influence of sex, age and symptomatology.

7.2.5 User of Beta-blockers

A special problem is the category of the non-diagnostic test result. A possible explanation for these patients not exercising sufficiently could be the use of beta-blocker medication. Cardiologists do not have a common policy regarding the use of beta-blockers during stress ECG testing. Of the 57 patients with coronary artery disease and on beta blocker medication, 30 (53%) did not reach the exercise criterion (Table 7.6). Of the remaining 244 patients with coronary artery disease who did not use beta-blockers, only 43 (18%) did not exercise sufficiently. The percentage of non-diagnostic tests was higher among the beta-blocker users and this difference was even greater in the patients without coronary artery disease, 73% (32 of the 44 patients using beta blockers) versus 23% (55 of the 235 pa-

tients not using beta blockers). As noted earlier, the use of beta-blocker medication may be the reason for patients failing to reach the required level of exercise. The use of beta-blockers may thus serve as an additional diagnostic factor in the interpretation of stress ECG results. Since the post-test probability for non-interpretable results is essentially the same for beta-blocker users (30/62=48%) and non-users (43/98=44%), beta-blocker use in these cases has little diagnostic value (Table 7.7). These findings suggest, thus, that the use of beta-blocker medication has definite, limiting effects on the diagnostic value of the stress ECG test, since an appreciable number of the stress ECGs performed are non-diagnostic.

Table 7.6 Relationship between use of beta-blocker medication and the stress ECG results in patients with (CAD+) and without significant coronary artery disease (CAD-)

	Number of patients		
	CAD+	Stress ECG non-diagnostic	Stress ECG -/+
BB+	57	30 (53%)	27
BB-	244	43 (18%)	201
	CAD-	stress ECG non-diagnostic	stress ECG -/+
BB+	44	32 (73%)	12
BB-	235	55 (23%)	180

BB+/- = whether or not beta-blocker medication was used during stress ECG testing;
 Stress ECG -/+ = negative/positive result stress ECG test

Table 7.7 Relationship between use of beta-blocker medication and a non diagnostic stress ECG result in patients with (CAD+) and without significant coronary artery disease (CAD-)

	Number of patients		
	CAD+	CAD-	total
Stress ECG, non-diagnostic			
BB+	30	32	62
BB-	43	55	98

BB+/- = whether or not beta-blocker medication was used during stress ECG testing

7.2.6 Conclusion

There is dependence between clinical variables and the test parameter (OR) and the usually presented crude values of a test parameter can be misleading. This dependence should be taken into account when estimating the post-test probability as it may well affect patient management.

Addendum 7.1 Estimations of the post-test probability (%) of stress ECG testing accounting for dependence (dep) and post-test probability assuming independence (indep)

Stress ECG:	Nil		< 1 mm		1 - 2 mm		> 2 mm		Non-diagn	
	Dep	Indep	Dep	Indep	Dep	Indep	Dep	Indep	Dep	Indep
Female < 50 years										
N.c.p./asymptomatic	6	3	5	4	14	9	37	31	7	6
Atypical angina	19	10	16	13	38	28	70	64	23	19
Typical angina	54	45	50	53	76	74	92	93	60	63
Female 50-59 years										
N.c.p./asymptomatic	12	6	10	8	26	18	57	50	15	12
Atypical angina	34	20	31	26	58	47	84	80	40	34
Typical angina	73	64	69	71	88	86	97	96	77	79
Female ≥ 60 years										
N.c.p./asymptomatic	17	10	15	14	35	28	67	65	21	20
Atypical angina	44	31	40	38	68	60	89	88	51	48
Typical angina	80	76	77	82	91	92	98	98	84	87
Male < 50 years										
N.c.p./asymptomatic	1	1	1	1	3	3	11	12	2	2
Atypical angina	5	3	4	5	11	11	32	36	6	7
Typical angina	36	36	32	44	59	66	85	90	42	54
Male 50-59 years										
N.c.p./asymptomatic	3	2	2	3	7	6	22	24	4	4
Atypical angina	10	7	9	10	23	21	52	56	12	14
Typical angina	36	36	32	44	59	67	85	90	42	54
Male ≥ 60 years										
N.c.p./asymptomatic	4	3	4	5	10	11	30	36	6	7
Atypical angina	14	12	12	16	31	33	62	70	18	23
Typical angina	46	50	59	69	69	78	89	94	52	69

8

The value of myocardial perfusion scintigraphy in the diagnosis of coronary artery disease

8.1 Literature review

Myocardial perfusion imaging with thallium-201 is a non-invasive method which can be used to evaluate the severity of coronary artery disease in patients with confirmed disease.

It can also be used to detect or to exclude coronary artery disease in subjects with (or without) symptoms. Only this second indication for its use will be assessed here. Thallium 201 myocardial perfusion imaging used as a tool in diagnosing patients with coronary artery disease is much more time-consuming and expensive than a stress ECG. High quality images are technically difficult to obtain and require considerable experience to be classified. The procedure also exposes the patient to radiation. General remarks concerning radiation have been made in Chapter 3.

A thallium scan shows whether the myocardium is normal, ischaemic or infarcted (Gould, 1984^{G6}; Soufer in Cohn P, 1985^{S6}; Iskandrian, 1985^{I3}; and Pohost, 1990^{P1}). The delivery of 201 Tl to the myocardial cell is a function of both regional blood flow and the fraction of the locally delivered tracer that is extracted by the myocardial cell. The defects in 201 Tl images are mainly due to regional abnormalities of blood flow. Most thallium scanning is done in combination with exercise. Its sensitivity in the detection of coronary artery disease is influenced by the amount of exercise performed and the increase in myocardial oxygen requirement achieved. The thallium tracer is injected just prior to completion of the exercise. It represents the initial distribution of 201 Tl.

Myocardial regions with reduced blood flow due to previous infarction or present ischaemia accumulate less thallium than normally perfused areas. Cells which have been infarcted will not take up thallium and ischaemic cells with a reduced extraction efficiency will take up less thallium than normal. Thallium distribution is not a static situation. Following initial distribution, thallium will continue to enter the myocardial cell from the large blood pool and thallium will also be cleared from the myocardium, the latter referred to as the "wash-out". The change from the initial myocardial distribution is called the redistribution. Scanning is repeated 2-4 hours after the initial injection. On this redistribution scan, thallium is present in both normal and ischaemic myocardium but not in areas of infarction. Hypoperfused areas accumulate less thallium and have a slower clearance (washout) of the thallium than normal myocardium. Areas of fibrosis have minimal perfusion and hence a low initial but persistent concentration of thallium. The site and the

extent of the defects on the scan correspond to the site and extent of the myocardial infarction and /or ischaemia. A scan without defects should only be called normal if the patient has exercised adequately, attaining a heart rate greater than 85% of the predicted maximum. A perfusion defect should be apparent in the same myocardial segment in more than one view.

The defects can be correlated with the coronary artery involved, left anterior descending, left circumflex or right coronary artery. Patterns can vary substantially due to normal variations in coronary anatomy.

In terms of qualitative image analysis, an image is either normal or it shows a perfusion defect which is reversible, partially reversible or irreversible (fixed). A patient may have different patterns in different myocardial segments or even within the same segment.

The reported sensitivity and specificity of exercise thallium scanning for detecting coronary artery disease without computer enhancement vary considerably, from 70 to 90% and 86 to 100%, respectively (Gould, 1984^{G6}; Iskandrian, 1985^{I3}). A high sensitivity is associated with a low specificity and vice versa, typical for an imperfect test (Table 8.1). Philbrick (1980)^{P4} and Niemeyer (1989)^{N5} discuss different causes for these findings.

Table 8.1 Ranges of sensitivity and specificity values for thallium scintigraphy, adapted from reported studies

Author	Year	Cases (n)	Sensitivity (%)	Specificity (%)
Turner ^{T5}	1978	66	68	97
Vogel ^{V1}	1979	42	72	94
Iskandrian ^{I4}	1981	194	79	95
Guiny ^{G7}	1981	35	88	88
Van der Wall ^{W4}	1988	391	84	92

Recently, Dilsizian advocated a two injection method, first used in the 1970's, which involved a re-injection of thallium after stress redistribution imaging, in order to enhance the detection of ischaemic but viable myocardium. This technique demonstrated a decrease of fixed defects, improving the detection of viable but jeopardised myocardium (Dilsizian, 1990^{D5}). Computer programs allowing quantitative analysis as developed by Watson (1981)^{W5} and Berger (1981)^{B11}, are now available. These allow analysis of 201 Tl myocardial kinetics, and the activity of thallium in myocardium can be quantified as well. Quantitative measurements of regional washout of thallium over 2 to 4 hours following injection can be obtained and compared to data from normal subjects.

The results of quantitative computer analysis as reported by Wackers (1986)^{W6} on the basis of 150 investigated patients were promising but he stated that quality control and attention to detail during acquisition and handling of the data were still essential for obtaining adequate results. He reported an improved sensitivity without loss of specificity; the sensitivity to detect significant coronary artery disease rose from 65% with qualitative analysis to 89% using quantitative analysis. The detection rate of single vessel disease also improved but less dramatically; however, false negative results still occurred more frequently in patients with single vessel disease.

Qualitative analysis is inaccurate for evaluation of the extent of coronary artery disease (Maddahi, 1981^{M5}; Ascoop, 1986^{A7}; Wackers, 1986^{W6} and Niemeyer, 1988^{N6}), especially in patients with multiple-vessel disease. Stenosed coronary arteries may cause a perfusion abnormality during stress that is not detected visually. This condition occurs in patients with multiple-vessel disease. Because the initial 201 Tl distribution reflects relative rather than absolute reduction in myocardial flow, areas with a lesser degree of hypoperfusion may appear relatively normal compared with the most severely hypoperfused segments. If there is equal obstruction of the coronary arteries (balanced disease) there may be an even flow in all areas and the scan may be reported as normal since no difference in relative thallium activity is observed. However, other signs may be present such as dilatation of the left ventricle and the presence of increased lung activity, indicating transient left ventricular failure due to diffuse ischaemia.

The relative concentration of thallium in the myocardium compared with that in the lungs has been correlated with the pulmonary capillary wedge pressure at the time of injection, and if the thallium concentration is equal, there is marked elevation of the wedge pressure. Increased lung uptake is due to prolonged pulmonary transit time and is an indication of left ventricular dysfunction.

Quantitative analysis also offers no advantage in the detection of multi-vessel disease (Niemeyer, 1989^{N5}).

Despite the advent of computer analysis of thallium scan images, visual interpretation is still commonly employed. Visual interpretation is open to inter- and intra-observer variability (Iskandrian, 1985^{I3} and Wackers, 1986^{W6}).

8.2 Results of thallium scintigraphy

Thallium scintigraphy was performed in 147 patients. In 115 patients both a stress ECG and a thallium scintigram study were performed. The thallium study followed the stress ECG test in 107 cases, a reverse sequence was used in 4 patients and both tests were carried out at the same time in 4 cases. In 32 patients, only a thallium test was conducted and in 14 of them an abnormal ECG at rest was the probable indication.

In 6 patients the exercise criterion was not reached and their results have been designated as non-diagnostic. As this finding is difficult to interpret and the number of patients is too low to handle them as a separate category in the analysis, these 6 patients are excluded in the following considerations. The number of 141 patients is a relatively small. The disease prevalence is also rather low (42%), suggesting that thallium scintigraphy was only performed in patients with distinct diagnostic characteristics. This selection bias impairs the external validity of the following results. One of the exclusion criteria for our study is the presence of a myocardial infarction. However, patients with a fixed defect on the thallium scintigram, which is suggestive of infarcted tissue, were not eliminated from the study. As discussed in Chapter 2 this test result was still recorded (as a positive test result) because with the present technique this finding is not specific enough for the presence of infarction (Pohost, 1990^{P1}).

There were 11 patients with a "fixed abnormality". Of these 11 patients, the left ventricular angiogram demonstrated normal contractility in 6, and one or more areas with hypokinesia were present in the remaining 5 patients. These findings justify the decision to include these patients.

The likelihood ratios for positive and negative scintigrams are given in Table 8.2.

Table 8.2 Results of thallium scintigram testing for patients with (CAD+) and without significant coronary artery disease (CAD-) and the likelihood ratio (LR) for the test results

	Number of patients		LR	95%-C.I.
	CAD+	CAD-		
Thal+	50 (84.7%)	62 (75.6%)	1.1	0.95-1.3
Thal-	9 (15.3%)	20 (24.6%)	0.63	0.31-1.3
	59 (100%)	82 (100%)		

95%-C.I. = 95% confidence interval of the LR; Thal+/- = positive or negative (and non-diagnostic) thallium scintigram

The crude LR for a positive test result is 1.2, with 0.52 for a negative result. These results seem rather disappointing as compared with other two non-invasive tests. The results have also been stratified for clinical variables (Table 8.3) when it becomes apparent that data are lacking for some strata. For patients with atypical angina and for females the LR for a negative test result is higher than the LR for a positive test result, indicating that the thallium scintigram does not discriminate at all for these categories.

Table 8.3 Thallium scintigraphy in patients with (CAD+) and without significant coronary artery disease (CAD-), and the likelihood ratio (LR), results stratified for sex, symptoms and age

	Number of patients		LR	95%-C.I.
	CAD+	CAD-		
Female				
Thal+	10	25	0.88	0.42-1.8
Thal-	5	8	1.38	0.45-4.2
Male				
Thal+	40	37	1.2	0.77-1.9
Thal-	4	12	0.37	0.12-1.2
N.c.p./asymptomatic				
Thal+	2	19		
Thal-	0	9		
Atypical angina				
Thal+	6	17	0.88	0.35-2.2
Thal-	4	8	1.3	0.38-4.2
Typical angina				
Thal+	42	26	1.0	0.61-1.6
Thal-	5	3	1.0	0.25-4.3
< 50 years				
Thal+	13	20		
Thal-	0	10		
50 - 59 years				
Thal+	18	28	1.1	0.59-1.9
Thal-	5	10	0.83	0.28-2.4
≥ 60 years				
Thal+	19	14		
Thal-	4	0		

95%-C.I. = 95% confidence interval of the LR; Thal+/- = positive or negative (and non-diagnostic) thallium scintigram. Where there are empty cells in the stratified cross tables, no LR's have been calculated

Furthermore, in Table 8.4 the OR as test parameter of thallium scintigraphy was determined with three different levels of stratification: no stratification at all (crude), stratified for a single clinical variable and (with help of logistic regression) stratified for all clinical variables. Due to the lack of data for some strata no ORs could be calculated for them. Thallium scintigraphy seems to discriminate between diseased and non-diseased patients only in males but not in females. Because a biological basis for this phenomenon seems improbable, this difference between males and females must be due to a different referral pattern for scintigraphy. The values of the OR of the remaining strata are all close to one, independent of the level of stratification. This suggests that thallium scintigraphy has, at least for this study population, a rather low discriminating value for diagnosing coronary artery disease.

Table 8.4 Crude odds ratio, odds ratio stratified for a single clinical variable and stratified odds ratio with correction for the influence of the two remaining clinical variables using a logistic regression model

	OR _{single}	OR _{logregr.}
Crude	2.3	
Stratified		
female	1.0	0.9
male	3.8	2.2
n.c.p./asymptomatic	?	1.4
atypical angina	0.9	1.4
typical angina	1.9	1.4
< 50 years	?	1.4
50-59 years	1.4	1.4
≥ 60 years	?	1.4

OR_{single} = stratified for a single variable; OR_{logregr.} = stratified for all variables with help of a logistic regression model, including only an interaction term of gender with scintigraphy

9

Coronary calcifications and coronary artery disease

9.1 Literature review

A number of studies have appeared indicating the value of (cine) fluoroscopy for detecting calcifications in coronary arteries as a non-invasive test for the evaluation of coronary artery disease (e.g. Bartel, 1974^{B8}; Hamby, 1974^{H5}; Aldrich, 1979^{A6}; Diamond, 1979^{D5}; Rifkin, 1979^{R7}; Margolis, 1980^{M6}; Hung, 1984^{H6}; Detrano, 1986^{D6}; Uretsky, 1988^{U1} and Locker^{L3}). Cardiac fluoroscopy as a method for assessing coronary artery disease is, however, still largely ignored. It has advantages over the stress ECG and thallium scintigram testing because its performance does not depend on the use of drugs or on the exercise capacity of the patient. Furthermore, it is a low cost, non-invasive test. (The presence of calcification in coronary arteries has also been reported as an incidental finding during conventional computed tomography as well as ultrafast computed tomography examinations of the thorax. However both these techniques are expensive and are certainly not low cost tests (Woodring, 1989^{W7}; Tanenbaum, 1989^{T6}).)

The major cause of coronary pathology in contemporary Western society is atherosclerosis. The structure and composition of the atherosclerotic plaque are well defined, but there is considerable debate concerning the genesis of the lesion. The intima is the cell layer principally involved in atherosclerosis. At least two different kinds of lesion may be found, the fibrous plaque and the complicated lesion. The complicated lesion is most commonly associated with arterial occlusive disease and consists of a fibrous plaque that has become altered due to haemorrhage, calcification, cell necrosis and mural thrombosis (Ross, 1976)^{R8}.

Post-mortem histological studies of excised coronary arteries in adults have demonstrated that discrete areas of calcification are almost invariably located in areas of advanced atherosclerotic disease. Small deposits may also be observed in a fibrous plaque associated with minimal luminal narrowing (Blankenhorn, 1961)^{B12}. Calcification in Monckeberg's medial sclerosis affects the media of the vessel exclusively and never results in luminal narrowing, in contrast to atherosclerosis. Radiographically, it may be difficult to distinguish between vascular calcification related to Monckeberg's medial sclerosis and that associated with atherosclerosis, although differentiation can always be made histologically (Blankenhorn, 1961)^{B12}. However the occurrence of Monckeberg's medial sclerosis in the coronary arteries is exceedingly rare (Blankenhorn, 1961^{B12}; Abrams, 1983^{A8}). Eggen (1965)^{E4}, Frink (1970)^{F4} and McCarthy (1973)^{M7} investigated the presence of calcification in coronary arteries using X-ray examinations of hearts obtained at

autopsy. The left coronary system showed a tendency for earlier and more extensive calcification than the right coronary artery. The proximal 3 cm of the major vessels were the prominent sites of calcification; there were no cases of calcification in more peripheral locations without proximal involvement. Frink found that the left anterior descending artery was the vessel most commonly involved. The presence of calcification at the site of the stenosis did not mean that the stenosis was always significant.

Snellen (1939)^{S7} reported on the fluoroscopic detection of calcification in the coronary arteries in his thesis, before the availability of image intensifiers. He stated that calcification was most frequently observed in the left coronary artery, partly due to a difference in occurrence rate and partly due to its easier detection. He preferred the left anterior oblique projection as in this projection calcifications in the anterior descending artery are superimposed and therefore more readily visualised. He advised fluoroscopy as an adjunct to other non-invasive methods in diagnosing coronary artery disease.

Lieber (1961)^{L2} correlated the fluoroscopic identification of coronary arterial calcification with clinical evidence of ischaemic heart disease. He stated that it proved to be difficult to determine whether the visualised calcification was present in the left main coronary artery or in the anterior descending branch. There was a higher prevalence of ischaemic heart disease in those patients with coronary calcification.

In a number of studies the (cine) fluoroscopic findings were correlated with the usual gold standard for diagnosing coronary artery disease, the coronary cine arteriogram. However, most studies in the literature reporting on (cine) fluoroscopy have not provided any useful information for our study. Nearly all published studies include patients with confirmed coronary artery disease - a previous myocardial infarction - and/or valvular heart disease, and therefore differ with regard to population composition (Bartel, 1974^{B8}; Hamby, 1974^{H5}; Aldrich, 1979^{A6}; Margolis, 1980^{M6}; Carboni, 1985^{C5}; and Uretsky, 1988^{U1}).

Patients with a previous myocardial infarction were, however excluded from the studies of Hung and Detrano (Hung, 1984^{H6}; Detrano, 1984^{D7}, 1986A^{D8}). Significant coronary artery disease was defined differently in the study of Hung making comparison invalid.

All the aforementioned reported studies are presented in an overview which examines population composition, technique used, definition of a significant lesion at arteriography which appears to be variable and often ambiguous (Chapter 3), etc., as well as values of some results (Table 9.1).

Bartel (1974)^{B8} investigated a population of 360 symptomatic patients. Further details are presented in Table 9.1. In the majority of the patients more than one vessel was calcified and the left anterior descending artery was the most frequently involved. The severity of the disease increased with the number of calcified vessels. There were 20 patients (15 males) under the age of 40 with calcification. All had significant disease but there are no details of the number of patients with significant disease who were less than 40 years old and had no calcifications. The number of patients with calcifications in the right coronary artery only was 14; 93% had significant disease. In a subgroup of 40 patients chosen at random Bartel also investigated whether the site of calcification found at fluoroscopy correlated with the site of the stenosis demonstrated at arteriography. There was a correlation in the majority; calcification was not at the exact site of the stenosis

in 10%. Bartel concluded that the finding of calcification in one or more vessels was highly suggestive of the presence of significant coronary artery disease.

Table 9.1 Overview of studies on coronary calcification

	Bartel (1974)	Hamby (1974)	Aldrich (1979)	Margolis (1980)	Hung (1984)	Carboni (1985)	Detrano (1986)	Uretsky (1988)
Cine fluoroscopy/Fluoroscopy	both	fl	fl	fl	fl	cfl	cfl	fl
Number of subjects	360	500	181	800	92	104	297	600
With/without symptoms?	with	both	both	with	with	with	with	both
Excluding myocardial infarction?	no	no	no	no	yes	no	yes	no
Sex	m/f	m/f	m/f	m/f	f	m/f	m/f	m/f
Age	all	30-70	21-55	all	29-65	34-68	mean 54	all
Significant lesion	70*	50*	50*	75d	70d	70d	50d	50d
Localisation of calcification - site stenosis?	yes	no	--	no	no	no	no	no
Sensitivity	55%	76%	65%	40%	78%	53%	66%	75%
Specificity	94%	78%	51%	93%	82%	93%	81%	78%
Likelihood ratio + test	9.1	3.4	1.3	5.7	4.3	7.5	3.4	3.4
	9.6-<50 year							
Likelihood ratio - test	0.4	0.3	0.6	0.6	0.3	0.5	0.4	0.3
	0.4-<50 year							
Prevalence of significant disease	74%	50%	53%	73%	30%	83%	46%	58%

fl: simple fluoroscopy; cfl: cine fluoroscopy - fluoroscopy recorded on cine film; *: percentage of stenosis, not defined whether diameter or area loss; ..d : % of stenosis as diameter

Hamby (1974)^{H5} investigated a population of 500 patients. Patients with more than one calcified vessel showed a more severe form of disease. The sensitivity and specificity of fluoroscopy for both sexes and older than 30 years were 76% and 78%, respectively. The likelihood ratio for a positive test was 3.4, and for a negative test 0.3. For ages below 50 years, sensitivity and specificity were 58% and 94%, respectively, with a LR+ of 9.6 and LR- of 0.4. Hamby concluded that the presence of calcification was of special importance in younger people more so than in older ones.

Rifkin (1979)^{R7} reviewed the literature on the importance of the finding of coronary calcification in the diagnosis of coronary artery disease. He stated that although calcification and stenosis were not synonymous, calcification always signified the presence of atherosclerosis. Fluoroscopically detectable calcification is, therefore, not an innocent finding even in the absence of major arterial narrowing. In such patients there may be a greater risk in the future for the development of obstruction.

Aldrich's (1979)^{A6} study population consisted of 181 patients with a high cholesterol level and with and without cardiac complaints. Coronary arteriography was performed in all subjects, including 93 asymptomatic cases who had a positive stress ECG and/or coronary calcification on fluoroscopy.

Margolis (1980)^{M6} reported on the diagnostic and prognostic importance of coronary calcifications. His population consisted of 800 subjects of both sexes. His study was an extension of the study reported by Bartel. He stratified for age but concluded that the results of the test were the same for 3 age groups, (< 45 years, 45-54 years, ≥ 55 years), which was in contrast to the findings of Hamby (1974)^{H6}.

Hung (1984) compared the results of three non-invasive tests; fluoroscopy, stress ECG and exercise thallium myocardial perfusion scintigraphy with coronary arteriography in 92 symptomatic females, suffering from typical angina, atypical angina

or non-specific chest pain. The (unstratified) likelihood ratio for a positive fluoroscopy test was 4.3 and 0.3 for a negative test. (The unstratified likelihood ratio for the stress ECG (positive: 1mm or more horizontal or downsloping ST segment depression) was 1.8 for a positive test result and 0.42 for a negative one. These values were adapted from the published report - 21 patients TP, 26 FP, 7 FN and 38 TN. For thallium scintigraphy the values were 8.3 for a positive test and 0.27 for a negative test result, - 21 TP, 6 FP, 7 FN, 58 TN.). The author concluded that cardiac fluoroscopy should be used more often.

Carboni (1985)^{C5} studied 104 symptomatic patients, males and females. Calcification was detected in approximately half the patients with significant disease. No calcification was demonstrated in subjects with normal vessels and it was found in only one with non-significant disease.

Detrano's study population consisted of 154 subjects (1984), later extended to 297 (1986B)^{D6}, including patients with cardiac symptoms or no symptoms. Patients with a history of or electrocardiographic evidence of myocardial infarction were excluded. Significant disease was defined as the presence of at least one lesion with a 50% diameter stenosis. Both sexes of all ages were studied.

In the 1984 study no values were reported for the test parameters of cine fluoroscopy, nor could these be extracted from the data in the published report. The sensitivity and specificity of the test were reported as 66% and 81% in the later 1986 study. (LR+ 3.4 and LR- 0.41). No stratified data could be extracted except for sex: for males LR+ = 4.3 and LR- = 0.36; for females LR+ = 2.5 and LR- = 0.52. The unstratified values of the likelihood ratio of stress ECG testing were LR+ = 2.3 and LR- = 0.47. for thallium scintigraphy LR+ = 3.4 and LR- = 0.34. The lower costs of fluoroscopy make it an inexpensive aid.

Uretsky (1988)^{U1} reported on fluoroscopy findings in 600 patients, including patients with valvular disease or a previous myocardial infarction. The intensity of the calcification was graded. The location was divided into single-, double- or triple-vessel involvement. Using LR's as a parameter of the fluoroscopy, an influence of age was detected. Both sexes demonstrated the same pattern. Symptomatology was not taken into consideration.

At greater age, more extensive disease was found, with more extensive calcification. Triple-vessel calcification, occurring in only 67 patients, had a high predictive value for the presence of coronary artery disease; only 2 false positive cases occurred in patients older than 65 years. Single-vessel calcification, usually in the left main - proximal left anterior descending artery, was most common but of no diagnostic value in patients older than 65. Even faint calcification increased the risk, especially in the younger patients less than 45 years of age. The author concluded that simple fluoroscopy could make an important contribution in the diagnostic process.

9.1.1 Conclusion

(Cine) fluoroscopy for the detection of calcification in the coronary arteries appears to be an easy to perform non-invasive test, with additional advantages of being independent of drugs and of the patient's exercise capacity, as well as having the benefit of relatively low cost. The majority of reported studies did not exclude patients with a previous infarction, which is necessary to determine its value as a non-invasive test for diagnosing coronary artery disease. Calcification is more common in the left coronary artery than in the right. Snellen (1939)^{S7} already reported that the left anterior oblique projection was the best projection

for detecting the calcifications. Differentiation between calcification in left main versus anterior descending artery is not readily possible upon screening without contrast medium in the vessels. There is discrepancy regarding the relationship between calcification and age and the relationship between calcification and sex is also not clear. Published values for the different parameters show variations and cannot be properly compared.

9.2 Results of (cine) fluoroscopy

9.2.1 Likelihood Ratio for coronary artery calcification

Out of a total of 778 patients examined calcification was present in one or more coronary arteries in 254 patients, 222 of whom had significant coronary artery disease (Table 9.2). Calcification was present at the site of a significant lesion in 176 of these 222 patients (80%). The remaining 46 patients showed only minor disease at the site of the calcification, but had at least one significant lesion elsewhere in their coronary arterial tree. 32 patients demonstrated calcifications but their vessels showed only minor stenoses or wall irregularity. No calcification was demonstrated in the coronary arteries of 524 patients but significant disease was apparent in 201 of them. The overall LR for a positive cine fluoroscopy test was 5.8 in our population and for a negative test 0.52 (Table 9.2). The 95% confidence intervals for the LRs were calculated using the formula for the approximate interval estimation of the risk ratio with cumulative incidence data (Rothman, 1986^R).

Table 9.2 Prevalence of coronary artery calcification in patients with (CAD+) and without significant coronary artery disease (CAD-) and the likelihood ratio (LR) for the test results

	Number of patients		LR	95%-C.I.
	CAD+	CAD-		
Calc-	201 (47.5%)	323 (90.9%)	0.52	0.47-0.58
Calc+	222 (52.4%)	32 (9.0%)	5.8	4.1 -8.2
- LCA	163 (38.5%)	30 (8.4%)	4.6	3.2 -6.6
- RCA/both	59 (13.9%)	2 (0.6%)	24.8	6.1 -101
	423 (100%)	355 (100%)		

95%-C.I. = 95% confidence interval of the LR; Calc+/- = whether or not calcification is present. The localisation of the calcification is specified by artery. LCA = left coronary artery only; RCA = right coronary artery only or both coronary arteries

According to the location of the calcification it was found that 193 patients showed calcification in only the left coronary artery. Only 6 patients had single-vessel calcification in the right coronary artery (4 with and 2 without significant disease) and 55 patients had calcification in both the left and right coronary arteries. The LR for a positive test in only the left coronary artery was 4.6 and for a positive test in the right coronary artery or in right and left coronary arteries was 24.8 (Table

9.2). The results of cine fluoroscopy with stratification for sex, age and symptomatology, with confidence intervals, are shown in Table 9.3. In order to reduce the number of categories without data or insufficient data, data have been merged. Patients with nonspecific chest pain have been combined with the asymptomatic patients. Patients have been grouped in 3 age categories: < 50 years; 50 - 59 years; \geq 60 years. In view of the small number of patients with calcification in only the right coronary artery or in both coronary arteries, these patients with calcification have been grouped into one single positive category.

9.2.2 Influence of clinical variables

As stated before (Chapter 4) to analyse the influence of the clinical variables age, sex and symptoms on the association between the test result and the disease probability and to detect the presence of collinearity, or collinearity combined with interaction, the (diagnostic) odds ratio (OR) was employed as test parameter and the data were analysed accordingly.

Stratification according to sex, age and symptoms demonstrates the presence of collinearity, as the stratified ORs differ from the crude OR (i.e. calculated for unstratified data), as well as interaction since the stratified ORs are different per stratum (Table 9.4). The crude OR is not the average of the stratified ORs, as the values stratified for symptoms demonstrate. This contradicts a statement made by Hlatky that the usual presented values of test parameters are average values (Hlatky, 1987^{H7}).

Table 9.3 Prevalence of coronary artery calcification in patients with (CAD+) and without significant coronary artery disease (CAD-), and the likelihood ratio (LR); results stratified for sex, symptoms and age

	Number of patients		LR	95%-C.I.
	CAD+	CAD-		
Female				
Calc+	53	12	7.6	4.3 -13.4
Calc-	39	146	0.45	0.36- 0.58
Male				
Calc+	169	20	5.0	3.3 - 7.7
Calc-	162	177	0.54	0.48- 0.61
N.c.p./asymptomatic				
Calc+	5	7	6.7	2.5 -18.1
Calc-	8	115	0.65	0.42- 1.00
Atypical angina				
Calc+	20	15	3.2	1.8 - 5.7
Calc-	33	111	0.70	0.56- 0.88
Typical angina				
Calc+	197	10	5.9	3.2 -10.7
Calc-	160	97	0.49	0.43- 0.56
< 50 years				
Calc+	35	6	9.7	4.3 -22.1
Calc-	49	134	0.60	0.50- 0.73
50 - 59 years				
Calc+	85	11	6.2	3.4 -11.2
Calc-	103	140	0.59	0.51- 0.67
\geq 60 years				
Calc+	102	15	2.9	1.9 - 4.5
Calc-	49	49	0.42	0.32- 0.55

95%-C.I. = 95% confidence interval of the LR; Calc+/- = whether or not calcification is present

Table 9.4 Crude odds ratio, odds ratio stratified for a single clinical variable and stratified odds ratio with correction for the influence of the two remaining clinical variables using a logistic regression model

	OR _{single}	OR _{logregr.}
Crude	11.1	
Stratified		
female	16.5	11.4
male	9.2	6.1
n.c.p./asymptomatic	10.3	9.7
atypical angina	4.5	4.6
typical angina	12.0	13.0
< 50 years	16.0	11.0
50-59 years	10.5	7.5
≥ 60 years	6.8	7.0

OR_{single} = stratified for a single variable; OR_{logregr.} = stratified for all variables with logistic regression

Furthermore, it can be seen that the OR for females is higher than that for males. This difference between the sexes could perhaps be explained by the influence of age or symptomatology, there being relatively more females in the lower age groups and in the n.c.p./asymptomatic group. Similarly, the differences between the three age groups may be explained by the influence of sex or symptoms, etc. An analysis correcting for all possible relationships between the three clinical variables with our limited number of patients is only possible with use of a logistic regression model. This method for estimating an OR been discussed in Chapter 4. The results obtained, now stratified for the three clinical variables, are shown in Table 9.4. Most of the values thus obtained differ considerably from the crude OR value (11.1). Furthermore, Table 9.4 shows that the ORs, stratified for only one clinical variable and the ORs stratified for all three variables differ in absolute values, but that the patterns are similar. In addition the different values for the ORs between the strata after correction for collinearity point to the presence of interaction.

In general a test is more discriminating and thus has a higher OR in patients with a higher extent of disease. This explains the high OR in patients with typical angina. However, the higher OR for females, younger patients and patients with n.c.p./asymptomatic, as compared with their complementary categories, can not be explained by this mechanism as these categories have less severe disease. The following observation may explain this discrepancy. As already explained in Chapter 4 the OR is the quotient of the products of the diagonal terms: $A \times D / (C \times B)$ as is shown in the following table.

	CAD+	CAD-
Fluoroscopy +	A	B
Fluoroscopy -	C	D

The OR of a specific stratum will be high if the term A or the term D is high relatively to the other terms. In patients with a high extent of disease the test result will likely be positive. Thus, the term A will be high and consequently the OR will be high. Patients without any coronary artery disease will never demonstrate calcifications and thus will have a negative fluoroscopy as well as a normal coronary arteriogram. Adding such patients to a stratum, for example by less stringent referral, increases the term D and consequently increases the OR also. This can be illustrated by Table 9.5.

Table 9.5 Percentage of patients with different manifestations of the atherosclerotic process and with the odds of having disease calculated for a positive and a negative cine fluoroscopy test; results stratified for gender, symptomatology and age

	Total	CAD–	CAD±	CAD+	Odds
Female					
fl+	100% (n= 65)	0%	18%	82%	4.4
fl–	100% (n=185)	58%	21%	21%	0.26
Male					
fl+	100% (n=189)	0%	11%	89%	8.5
fl–	100% (n=339)	35%	17%	48%	0.91
N.c.p./asymptomatic					
fl+	100% (n= 12)	0%	58%	42%	0.71
fl–	100% (n=123)	67%	27%	6%	0.06
Atypical angina					
fl+	100% (n= 35)	0%	43%	57%	1.3
fl–	100% (n=144)	58%	19%	23%	0.29
Typical angina					
fl+	100% (n=207)	0%	5%	95%	19.7
fl–	100% (n=257)	25%	13%	62%	1.6
< 50 years					
fl+	100% (n= 41)	0%	15%	85%	5.8
fl–	100% (n=183)	57%	16%	27%	0.36
50-59 years					
fl+	100% (n= 96)	0%	11%	89%	7.7
fl–	100% (n=243)	42%	16%	42%	0.73
≥ 60 years					
fl+	100% (n=117)	0%	13%	87%	6.8
fl–	100% (n= 98)	23%	27%	50%	1.00

fl+/- = test result of cine fluoroscopy; CAD+ = significant coronary artery disease; CAD± = coronary artery disease but no lesion of haemodynamic significance; CAD– = no signs of coronary artery disease; Odds = Odds of having significant disease

In this table patients with no significant coronary artery disease are divided into patients with radiologically complete normal vessels and patients with only minimal disease in the coronary arteries. By definition, patients with normal vessels can not have calcifications in the wall and thus these patients contribute only to the term D. This is the case for females compared with males (58% vs. 35%), for patients younger than 50 compared with patients >60 years old (57% vs. 23%) and for patients with n.c.p./asymptomatic compared with patients with typical angina (67% vs. 25%). Term A consists of patients with significant disease as well as calcification(s). For gender and age there are only minimal differences between the percentages of the distinct strata. In contrast these percentages increase markedly when the symptoms become more typical (with increasing severity: 42%, 57% and 95%). Thus, the pattern of the ORs for the different strata can be

explained by these two processes. As discussed in section 9.1, most studies in the literature reporting on (cine) fluoroscopy provide no useful information for our study. Nearly all investigations include patients known to have coronary artery disease - a previous myocardial infarction - and/or valvular heart disease, and therefore differ with regard to population composition. Patients with a previous myocardial infarction were however excluded in the studies of Hung and Detrano (Hung, 1984^{H6}; Detrano, 1986^{D6}), but significant coronary artery disease was defined differently in these two studies, again making comparison invalid.

9.2.3 Probability analysis

The presence of collinearity and interaction imply that the assumption of independence is not correct. The usually recommended method of Bayes for calculating the post-test probability would involve the use of conditional LR_s (or sensitivity and specificity) which is not practical for the number of clinical variables included in this analysis. Therefore, a logistic regression model was used here also to estimate the post-test probability for all combinations of sex, age and symptom groups. The discriminating value of (cine) fluoroscopy is graphically displayed in Figure 9.1. As judged by the post-test probabilities the test separates patients properly with respect to disease status. It is in the probability range 10-20% where the decision has to be made whether or not to perform another (more expensive) non-invasive test or even an arteriogram. If only (cine) fluoroscopy is used as non-invasive test, the test result could influence the decision for referral for an arteriogram for male and female patients with n.c.p./asymptomatic and for females with atypical angina (Figure 9.1). (Furthermore, patients in all other categories could be referred for arteriography, independent of the result of the (cine) fluoroscopy.) As discussed before, differences between estimation of the post-test probability by logistic regression and estimation without taking dependence into account may be of importance as differences may well lead to different management of the patient. Differences are present in all categories (addendum 9.1.) and will increase further if the diagnostic chain (process) is made longer. The differences, as demonstrated in Figure 9.2 for the 10-20% post-test probability range, may indeed influence the further management of the patients.

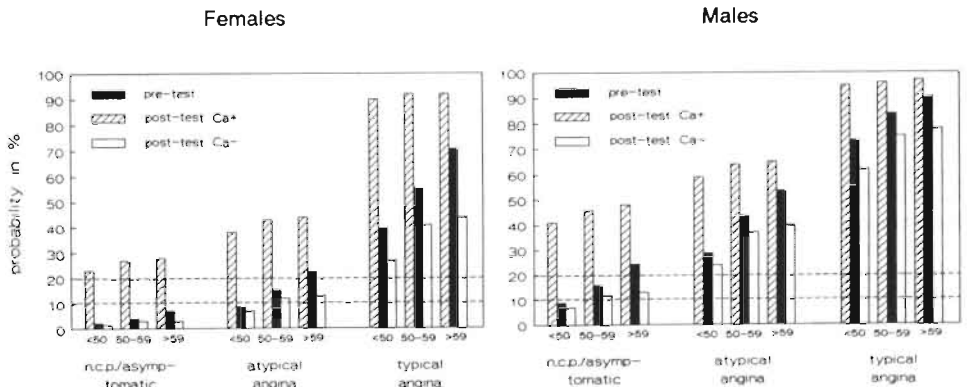


Figure 9.1 Pretest and posttest probabilities for females (left) and males (right) and for different age and symptom categories. The dotted lines mark the probability range 10%-20% where the decision has to be made whether or not to perform an arteriogram

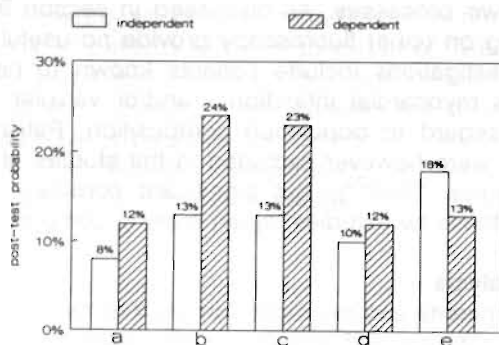


Figure 9.2 Some examples of differences between the post-test probability assuming independence and the post-test probability corrected for collinearity and interaction, which may lead to differences in patient management.

(a) males, 50 - 60 years, n.c.p./asymptomatic, cine fluoroscopy negative; (b) males, < 50 years, atypical angina, cine fluoroscopy negative; (c) females, < 50 years, n.c.p./asymptomatic, cine fluoroscopy positive; (d) females, 50 - 60 years, atypical angina, cine fluoroscopy negative; (e) females, \geq 60 years, atypical angina, cine fluoroscopy negative

9.2.4 Conclusion

(Cine) fluoroscopy appears to be a useful non-invasive test for diagnosing coronary artery disease as judged by its discriminating effect, especially for patients of either sex with n.c.p./asymptomatic, and females with atypical angina; all patients with a low pre-test probability. There is a dependence between clinical variables and the test parameter (OR), and the usually presented crude values of a test parameter can be misleading. This dependence has two components: collinearity and interaction. To correct for this dependence for more than one clinical variable, the data should be analysed with a logistic regression model. In this way, the analysis of the test performance is adjusted for collinearity and accounted for interaction. This dependence should be taken into account when estimating the post-test probability as it may well affect patient management.

Addendum 9.1 Estimations of the post-test probability (%) of cine fluoroscopy accounting for collinearity and interaction, and post-test probability assuming independence shown in brackets

Symptoms/Age	cine fluoroscopy +		cine fluoroscopy -	
	Male	Female	Male	Female
N.c.p./asymptomatic				
< 50 years	41 (31)	23 (13)	7 (4)	2 (1)
50 - 60 years	46 (48)	27 (24)	12 (8)	3 (3)
\geq 60 years	48 (63)	28 (38)	13 (13)	3 (5)
Atypical angina				
< 50 years	59 (63)	38 (38)	24 (13)	7 (5)
50 - 60 years	64 (78)	43 (56)	37 (24)	12 (10)
\geq 60 years	65 (87)	44 (70)	40 (38)	13 (18)
Typical angina				
< 50 years	95 (96)	90 (83)	62 (55)	27 (30)
50 - 60 years	96 (97)	92 (91)	75 (72)	41 (47)
\geq 60 years	97 (98)	92 (95)	78 (83)	44 (63)

10 | Literature review of proposed protocols: the need for additional information

10.1 Literature review

In the previous chapters the values of different diagnostic variables, if used alone, have been discussed. This chapter reviews published studies of proposed protocols involving one or more of the different available non-invasive tests for diagnosing coronary artery disease (Diamond, 1979^{D5}, 1980^{D9}, 1983^{D10}, Hung, 1984^{H6}, Detrano, 1984^{D7}; Melin, 1985^{M8}, Dunn, 1982^{D11}, Patterson, 1982^{P2}, 1984^{P8}, Weintraub, 1984^{W8}; Chaitman, 1984^{C6}).

These studies are based on the concept that post-test probability levels of < 10% or > 90% (or 20% or 80%) are adequate end point criteria for diagnostic testing, "establishing" the absence or presence of disease.

Bayes' theorem is used to estimate the post-test probability. If more tests (signs or symptoms) are conducted in sequence the post-test probability value of the first test is the pre-test probability value of the following, second test. Independence is assumed between the different non-invasive tests.

Diamond (1979)^{D5} reported on (a) values for the probability of having coronary artery disease, based on age, sex and symptomatology (already discussed in Chapter 5) and (b) an assessment of 3 non-invasive tests, stress ECG testing, thallium scintigraphy testing and fluoroscopy (to detect calcium deposits in coronary arteries), by taking into account their value in helping to estimate the post-test probability of having coronary artery disease. In this probability analysis involving these non-invasive tests he applied sensitivity and specificity values which he had obtained by pooling the values taken from a number of publications. He used those as published by Bartel (1974)^{B8} for the fluoroscopy. He further re-interpreted the stress ECG tests results not merely as positive or negative but in a semi-quantitative way; each degree of exercise-induced ST segment response with its own specific sensitivity and specificity values. It is not clear from the study if patients with a previous myocardial infarction were excluded.

He finally concluded that the use of non-invasive tests could be rationalised by applying probability analysis. Diamond's prospective study of 1980, a continuation of his 1979 study, involved 43 infarction-free patients, all of whom were subjected to arteriography. A stenosis with a diameter loss of 50% was considered a significant lesion.

All patients were subjected to the 3 non-invasive tests enabling post-test probability estimates of all possible test combinations. Any degree of ST segment depression was called positive, and assigned to a category, <0.5mm, 0.5-0.9mm, etc.

He introduced in this article the concept that post-test probability levels of $<10\%$ and $>90\%$ were adequate end-point criteria for diagnostic testing, "establishing" the absence or presence of disease. The test combination of fluoroscopy and stress ECG would have correctly placed 26 of the 43 patients in one of the two categories, $<10\%$ or $>90\%$. A subsequent thallium scintigram in the 17 remaining patients, placed initially in the indeterminate group, could have increased the number of subjects in the two categories $<10\%$ or $>90\%$ by another 8, thus a total of 34 correctly categorised patients.

The test combination of only a stress ECG and a thallium scintigram would have only classified 27 patients correctly and would have increased the total costs of testing, as compared with the more selective use of thallium scintigraphy after fluoroscopic and stress ECG testing as described above. The protocol thus suggested is depicted in Figure 10.1.

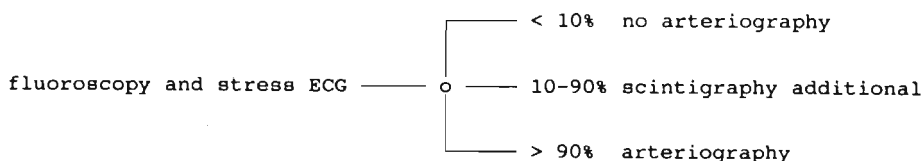


Figure 10.1 Decision algorithm, for all patients regardless of pre-test probability (from Diamond, 1980)

Thus, the value of the probability estimate based on the clinical variables (the pre-test probability) and its eventual help in deciding to refrain from non-invasive testing and/or arteriography was not determined in this publication. Diamond's study of 1983 consisted of 1097 patients of whom only 170 underwent coronary arteriography following the non-invasive tests. Each of the subjects underwent stress ECG and subsequently at least one additional non-invasive diagnostic test, cardiac fluoroscopy, thallium perfusion scintigraphy, cardio-kymography or technetium-gated blood pool scintigraphy. Patients with a previous myocardial infarction were excluded.

Each patient underwent from 2 to 5 non-invasive diagnostic tests, (average of 3 tests per patient). The angiographic studies were considered to show significant disease if there was a lesion of at least 75% loss in cross-section area in a major vessel.

A commercially available software program called CADENZA, which could be used on a microcomputer and which estimated the post-test probability based on Bayes' theorem but did not account for dependence, was employed to analyze the results of the various clinical variables and non-invasive tests. The various combinations of non-invasive tests resulted in a total of 2180 different combinations and probability estimates for the 170 patients who had an arteriogram.

Age, sex and symptoms proved to be important discriminators for coronary artery disease. This finding emphasises the importance of the clinical history as a potent diagnostic test, making it possible to classify patients on this basis into subgroups with a very high or very low probability of having coronary artery disease.

The presence of multi-vessel disease increased as a monotonic function of post-test probability. With a post-test probability below 25%, single-vessel disease was slightly more common than multi-vessel disease. Above a probability of 75%, multi-vessel disease predominated.

Diamond concluded that his data indicate that Bayes' theorem in general and the CADENZA program in particular constitute accurate, clinically applicable methods for quantifying the prevalence of angiographic coronary artery disease and the risk of multivessel disease.

Hung (1984)^{H6} reported on the choice of the three non-invasive tests in a group of 92 women without a previous myocardial infarction. The selected criteria for a positive test differed from the 2 above-mentioned studies. A stenosis of 70% of the arterial diameter was considered significant disease. The stress ECG test was deemed positive if there was an ST segment depression of 1 mm or more. Significant disease was present in 60% of these females with typical angina, in 30% with atypical angina and in 7% with non-specific chest pain.

Cardiac fluoroscopy or thallium scintigraphy provided more diagnostic information than stress ECG in this group of females over a wide range of clinical subsets. However, the study population was small and the influence of age was not assessed. Fluoroscopy seemed more cost effective.

Hung devised different strategies of non-invasive testing for subgroups varying in probability based on the clinical variables age and symptoms, (Figure 10.2). No further tests were recommended for females with non-specific chest pain.

Thus, the author discussed (on the basis of the magnitude of change in value of pre- and post-test probability) which test or combination should be performed. Advice on whether or not to perform arteriography on the basis of post-test probability values obtained was not offered.

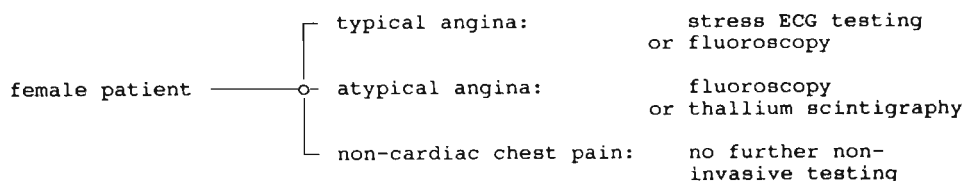


Figure 10.2 Decision algorithm, for females, (from Hung, 1984)

Thus, this approach related the use of the non-invasive tests to the symptomatology of the (female) patient; it reduced the number of non-invasive tests required to a single one. This protocol differed from previous protocols in that the non-invasive test to be used varied for the different categories of patients.

The following study by Detrano expands this new concept, as judged by the protocol finally proposed. Detrano's study (1984)^{D7} included 154 subjects, all of whom were subjected to the three non-invasive tests, with arteriography as the conclusive investigation. Patients with a previous myocardial infarction were excluded. The probability values used were based on sex, age and symptoms as published by Diamond (1979)^{D5}. The use of different decision algorithms and comparison was introduced in this article. Values of <20% and >80% were proposed as adequate end-point criteria for diagnostic testing in this study.

- (1) Detrano compared 2 strategies based on the principle:
 "-All subjects with a disease probability of at least 20% will undergo coronary arteriography". The results of applying this rule without any non-invasive testing and after non-invasive testing were compared (Figure 10.3).

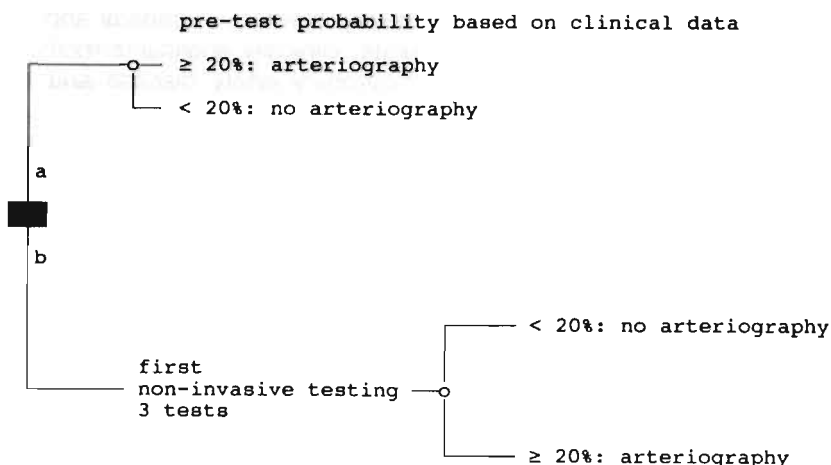


Figure 10.3 Comparison of two decision algorithms, (from Detrano, 1984)

Out of the 154 patients involved, each of the two paths resulted in a patient with significant disease who nevertheless were assessed in the "less than 20%" category. Neither patient underwent an operation, however. The number of arteriograms without significant disease in the $\geq 20\%$ category dropped from 56 to 32 in the second pathway. A further reduction in the number of non-invasive tests required is possible, as will be seen in the following paragraph.

- (2) Detrano further tested the following protocol, which would be less expensive compared with the previous ones (due to a reduction in non-invasive testing):

"-All subjects with a pre-test probability of at least 80%, based on clinical variables, will undergo coronary arteriography without non-invasive testing.

-All subjects with a pre-test probability of less than 20% will undergo neither non-invasive testing nor arteriography.

-Subjects with a pre-test disease probability between 20 and 80% will be tested non-invasively and coronary arteriography will be performed only if the post-test probability is $\geq 20\%$ " (Figure 10.4).

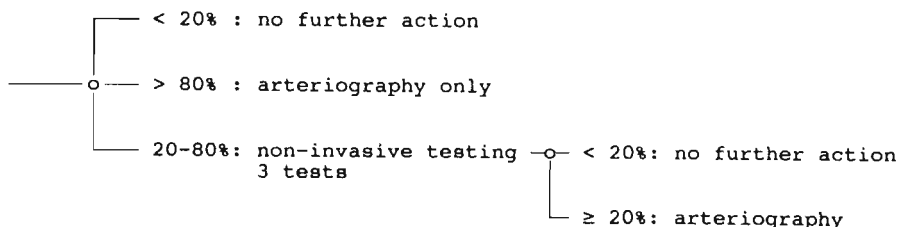


Figure 10.4 Decision algorithm, (from Detrano, 1984)

This strategy resulted in a further reduction in the number of subjects submitted to non-invasive testing as compared with the algorithm presented in the previous Figure 10.3.

Conclusion

The protocol as proposed by Detrano (Figure 10.4), involving 3 non-invasive tests which are used in a certain subset of patients, seems to be the best protocol as reported in the literature. The margins set (20% and 80%) are, however, arbitrary.

A number of studies have examined the combination of two non-invasive tests: the stress ECG and thallium scintigraphy, with arteriography as the gold standard (Dunn, 1982^{D11}; Weintraub, 1984^{W8}; Melin, 1985^{M8}); stress ECG testing combined with fluoroscopy (Rifkin, 1979^{R7}); and thallium scintigraphy combined with fluoroscopy (Chaitman, 1984^{C6}). As these studies only involve two tests they are of less importance for the aim of the present investigation.

Summary

Three non-invasive tests, the stress ECG, thallium scintigraphy and (cine)fluoroscopy for coronary calcification, are all imperfect tests for diagnosing coronary artery disease. The second one is much more expensive than the other two. It is also believed that the third one should be used more often; further evaluation seems justified. Patients with a previous myocardial infarction, and thus known coronary artery disease, should be excluded in such an evaluation.

It seems reasonable to conclude at this stage that coronary arteriography should not be performed in patients with a low probability of having coronary artery disease. The same may perhaps apply for the non-invasive testing. An arteriogram as the only investigation seems justified in subjects with a high probability. Patients could be classified into subgroups of low, intermediate and high probability with 10 and 90% or 20 and 80% as set margins, as attempted in some reported studies.

Providing a basis for these arbitrarily chosen margins for low, intermediate and high risk would be an asset.

10.2 Limitations of proposed protocols and need for additional information

A number of protocols have been described in the literature. Consideration of the chance of having coronary artery disease, thus using a probability analysis, forms the core of these protocols. Patients are categorised according to their post-test probability, while the margins of the categories are arbitrarily set. The decision whether to perform non invasive tests or arteriography depends on the category to which the patient is assigned. There is no differentiation in the use of the various non invasive tests in the published protocols. Thus, for instance all patients with a post-test probability, based on clinical diagnostic variables, below a certain percentage are not entitled to an arteriogram or further non-invasive testing. Patients with a probability percentage in a certain range will all undergo a stress ECG, etc; instead of for instance males, > 60 years with typical angina will have a stress ECG only; females, 50-60 years with atypical angina, fluoroscopy and if negative only a stress ECG. Comparison of the costs of the various non invasive tests is also not taken into consideration. There are therefore reasons to design our own protocol.

As has been demonstrated, there is a wide variation in the values of the different parameters. Apart from the cause, these variations impair the choice of which value to use for the probability estimations in making one's own protocol. Many publications include patients with confirmed coronary artery disease, for example with a previous myocardial infarction. Data on such patients are not relevant for the aim of our study, as our protocol is only intended for patients suspected, to a varying degree, of having coronary artery disease. This is another disadvantage of using results based on such data.

Independence versus dependence of test results

Objections have been raised against the method used to calculate the post-test probabilities. Various authors have indicated that the presence of dependence between tests invalidates the use of Bayes' theorem while assuming independence.

Using the results as reported in a study by Patterson (1982)^{P2}, Weintraub (1984)^{W8} found some dependence between stress ECG testing and thallium scintigraphy testing, especially for the sensitivity. He called this dependence for sensitivity "partial positive test interaction", physiologically explained by the concept that a perfusion defect detected by thallium scintigraphy represents inadequate flow leading to ischaemia, whereas the ST segment depression in the stress ECG represents the same ischaemia in that zone. Both tests are therefore partially dependent. He discussed the idea that the application of Bayes' theorem (while assuming independence) would be appropriate if the tests were independent of each other, that is, if the tests evaluated unrelated pathophysiological variables. From a statistical point of view this would mean that the sensitivity and specificity of one test would be the same whether the other test was positive or negative. However, Weintraub just concluded that the application of Bayes' theorem with the assumption of (and assuming) independence was still clinically useful despite this evidence of test dependence.

Detrano (1984)^{D7} in his article on the use of different non invasive tests for diagnosing coronary artery disease and using Bayes' theorem, discussed the issue of dependence, and noted dependence between a number of the non-invasive diagnostic tests used but voiced the opinion that just calculating the post-test probability and ignoring these interactions did not invalidate the end result. A proper comparison, however, was not performed.

Also Sackett's calculations (using Diamond's 1979 data) suggested that stress ECG testing and thallium scintigraphy testing were not independent, and he called this dependence "concordance", leading to some (6%) overestimation of the post-test probability (Sackett, 1985^{S1}). Furthermore, he concluded without any argument, that this was not a problem with short sequences of two or three diagnostic tests.

McNeil calculated the post-test probability assuming independence and also accounting for dependence between two tests used to diagnose reno-vascular disease, the intravenous urogram and the renogram, with the arteriogram as gold standard (McNeil, 1975^{M9}). If both tests were positive the post-test probability was 95% in the first instance and 74% when dependence was taken into account (Chapter 4). A positive test loses some of its information value if known that a second test is also positive. An estimate of the post-test probability when both tests are positive, but assuming independence, will be too high. A conditioned LR should be used as explained in Chapter 4.

In recent clinical reports both Macartney (1987)^{M10} and Knill-Jones (1987)^{K5} discussed the problem of over-estimation of the probability of disease if no adjustment is made for dependence and referred to methods such as logistic regression in order to modify these calculations.

Fletcher stated in the second edition (1988)^{F1} of his book *Clinical Epidemiology* (page 69) that it was unlikely that multiple tests for most diseases were truly independent of one another. If the assumption that tests are completely independent is wrong, calculation of the probability of disease from several tests would lead to a misleading post-test probability value. A solution to the problem was not offered.

However, Diamond (1980)^{D9} argued, regarding only the physical concepts of the tests, that stress ECG and thallium scintigraphy testing were independent, pointing out that the first is an electrophysiological and the second a hydrodynamic event. This would also apply to fluoroscopy, the screening for calcification in coronary arteries, a metabolic phenomenon.

Furthermore, there has been a heated debate in current cardiology literature in the USA, especially between Hlatky and Diamond, as to whether the test parameters such as sensitivity, specificity, or the likelihood ratios, are constant, i.e. whether or not they vary with respect to conventional clinical characteristics and change only with thresholds which are arbitrarily chosen as definitions of disease and abnormality (Hlatky, 1984,^{H8} 1986^{H9} and 1987^{H7}; Diamond, 1986^{D13,D12,D14}; Tavel, 1987^{T7}).

This issue of (in)dependence is analogous to the issue of (in)dependence between the results of different tests as discussed in the previous section. The influence of clinical variables on test performance was discussed by Hlatky in 1984. He analysed data related to stress ECG testing and formed the opinion that clinical factors such as age, sex and symptoms, affect the sensitivity and specificity of the stress ECG test. The type of chest pain was also reported to affect the sensitivity of the stress ECG. He voiced the opinion that it makes little sense to assume that a test is likely to have an equal probability of being positive in all patients with disease. That would ignore the fact of the spectrum of disease and it is more likely that a test will be positive if the extent of the disease is more pronounced. The extent of disease will vary with age, and also with symptoms (e.g. unstable angina versus atypical angina) and may also differ for sex. This influence of the extent of disease was not initially supported by Diamond until 1986 when he agreed that the extent of disease has an influence on test results (Diamond, 1986)^{D13}. Hlatky furthermore concluded that these clinical variables influence the test parameters and that this is in addition to (and independent of) the degree of coronary artery disease (Hlatky, 1984^{H8}). Thus, Hlatky emphasised the existence of test dependency between symptomatology (one "test") and the stress ECG (a subsequent test). Not taking the dependence into consideration would falsely assume that one could use the same likelihood ratio value of the stress ECG test result in a 22-year old female patient with atypical angina and in a 70-year old male with typical angina and the same test result.

Both Douglas and Eaker also believed that the sexes differed in cardiac diagnostic testing, thus indicating the dependence between a clinical variable and the outcome of a test (Douglas, 1986^{D15}; Eaker, 1988^{E5}). This possible influence of clinical variables on the test performance implies an additional reason why the usually recommended method of Bayes, while assuming independence between

diagnostic features such as age, sex, symptoms and a performed diagnostic test, ought not be used to calculate the post-test probability.

Making allowance for dependence becomes even more important if the diagnostic chain is longer, i.e. if more clinical variables and more non-invasive tests are involved in the diagnostic work-up (Sackett, 1985^{S1}; Fletcher, 1988^{F1}).

In a probability analysis, taking this influence of the clinical variables into account, one needs information about sex, age and symptoms as well as the results of the different tests in each individual patient. The data provided in studies published to date are insufficient and are of no use for our aim.

We have therefore chosen to formulate our own protocol based on findings derived from our own study population, patients seen at the De Wever Hospital.

Both confounding and interaction between clinical variables and the test performance occur, as demonstrated for stress ECG in Chapter 7 and cine fluoroscopy in Chapter 9. This implies that the use of the usually recommended method of Bayes while assuming independence is probably not appropriate for calculating the post-test probability, as it may very well be inaccurate.

As elaborated in Chapter 5 it could be decided that coronary arteriography is only indicated if the probability of disease is estimated at 15% or more. Consequently, to construct the diagnostic flow charts for the different categories of patients it is necessary to estimate this probability at all possible steps of the diagnostic process. This information has to be deduced from the data on clinical variables, results of the various non-invasive tests and coronary arteriographic findings which are recorded for each of the patients in the study population. As described, six diagnostic variables have been collected from each of our patients: data regarding symptomatology, age, sex, as well as the results of three non-invasive tests if performed. The data have been sorted as follows:

Symptomatology has been sorted into three groups:

- a) typical angina
- b) atypical angina
- c) nonspecific chest pain combined with the asymptomatic, as defined previously.

Age into three groups:

- a) 49 and less
- b) 50 - 59
- c) 60 and older.

Sex into male and female

Stress ECG into 4 categories:

- a) negative combined with ST segment depression $< 1\text{ mm}$
- b) ST segment depression of $\geq 1\text{ mm}$ but $< 2\text{ mm}$
- c) ST segment depression of $\geq 2\text{ mm}$
- d) non-diagnostic result, (as defined before in Chapter 2)

Thallium scintigram into 2 groups:

- a) positive
- b) negative.

Cine fluoroscopy for calcification into positive or negative.

As discussed in Chapter 10 this information can be deduced with and without accounting for dependence between the different diagnostic variables. If independence between these diagnostic variables is assumed, the post-test probabilities can be estimated by using the crude LRs as calculated in the previous chapters: for stress ECG testing (Chapter 7), for thallium scintigraphy (Chapter 8) and for fluoroscopy (fl) (Chapter 9).

Estimation of the post-test probability while accounting for collinearity is in theory possible with use of stratified LRs, but it demands large numbers of patients in order to provide necessary data for all the strata. However, with help of a logistic regression model, as described in Chapter 4, it is possible to give an estimate of the post-test probability while accounting for collinearity and taking interaction into consideration. In the first part of this chapter (sections 11.1 and 11.2) the results of these two approaches will be compared. In the second part, after the selection of the most appropriate method, the different steps in the construction of the diagnostic flow chart will be described.

11.1 Assuming independence between diagnostic variables

The easiest approach is to assume independence between all diagnostic variables. The post-test probability after non-invasive testing can be calculated using the crude LRs (Table 11.1) of the non-invasive tests as reported in previous chapters for stress ECG testing (7), thallium scintigraphy (8) and fluoroscopy (9).

Table 11.1 LRs (crude) values for the three non-invasive tests

Test	LR
Fluoroscopy	
positive	5.8
negative	0.52
Stress ECG	
≥ 2 mm	6.02
1 - < 2 mm	1.29
< 1 mm	0.51
non-diagnostic	0.78
nil	0.37
Thallium scintigraphy	
positive	1.2
negative	0.52

In the following this method of estimating post-test probabilities is demonstrated for males, < 50 years, with atypical angina undergoing fluoroscopy, stress ECG and thallium scintigraphy sequentially.

Based on the clinical variables sex, age and symptoms, and assuming independence between these clinical variables, the probability of having coronary artery disease for this category in our study population is 23%. Thus, the pre-test odds are 0.29 ($0.23/(1-0.23)$). If, for example, the fluoroscopy is positive, which has a LR of 5.8, the post-test odds for this test results are 1.7 (0.29×5.8) and consequently the post-test probability is 63% ($1.7/(1.7+1)$).

Thus, after fluoroscopy and before stress ECG testing, the odds for the patient with the above test result are 1.7. If, for example, the result of the stress ECG is negative, the LR is 0.37 and the post-test odds will be 0.68 (1.7×0.37) and the post-test probability is 41% ($0.68 / (0.68 + 1)$).

In an identical way the post-test probabilities can be calculated using thallium scintigraphy. The complete tree of the post-test probabilities for this category with all possible test results is shown in Figure 11.1.

This kind of diagnostic flow chart can be constructed in an identical way for patients with other clinical characteristics.

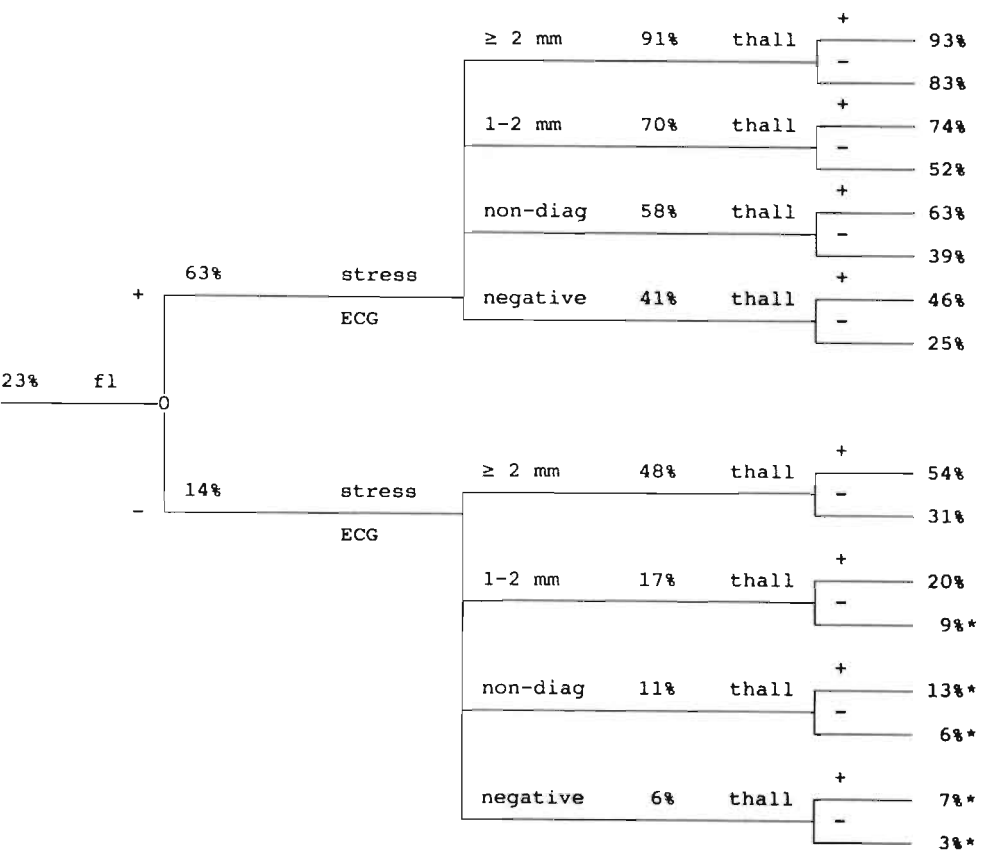


Figure 11.1 Probability estimates of having disease, while assuming independence, for male, < 50 years, atypical angina and sequential use of fluoroscopy (fl), stress ECG and thallium scintigraphy (thall).

* No arteriography;

23% = pre-test probability based on sex, age and symptoms, while assuming independence between these clinical variables

11.2 Accounting for dependence between diagnostic variables

If one wishes to take into account dependence between the different diagnostic variables, a more sophisticated way, i.e. a logistic regression method (Chapter 4), has to be used. With help of this method, the probability of having coronary artery

disease based on the clinical variables can be estimated, as well as the post-test probabilities of non-invasive testing. The values for probabilities after testing for the same category of patients as in the preceding, i.e. males, < 50 years with atypical angina, can be determined and are shown in Figure 11.2.

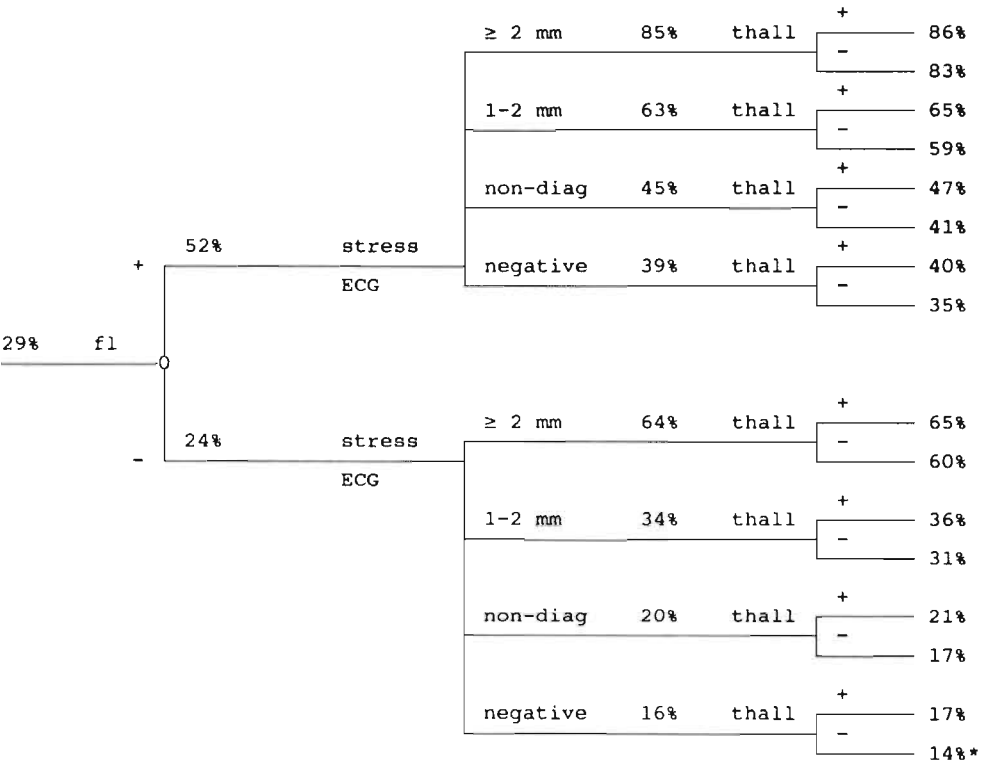


Figure 11.2 Probability estimates of having disease, while accounting for dependence, for male, < 50 years, atypical angina and sequential use of fluoroscopy, stress ECG and thallium scintigraphy.
 * No arteriography
 29% = pre-test probability based on sex, age and symptoms, while accounting for dependence between these clinical variables. (This value is higher than the value assuming independence between the clinical variables. This is not unexpected as has been explained previously.)

Comparing Figure 11.1 with Figure 11.2 there appear to be substantial differences in the post-test probabilities between the independent method and the method accounting for dependence. This is not unexpected as has been explained previously. However, using a logistic regression model to estimate post-test probabilities is more elaborate but it is more justified if the differences in results of the two methods also lead to differences in management of the patients. As has been discussed in Chapter 5, the management of the patient is determined by the value of $P_{\text{threshold}}$, which has been set at 15%. Thus, arteriography will be advised if the post-test probability is $\geq 15\%$, and the patient will be denied an arteriogram if the post-test probability is $< 15\%$.

Again comparing the two figures, one can see that the two different methods do lead to differences in patient management. Patients with a non-diagnostic or negative result of the stress ECG will be withheld arteriography if the independent method is used, but will be referred for arteriography if the other method is used. Furthermore, there are differences in the referring strategy after thallium scintigraphy. This effect of the method of analysis on patient management is also applicable to patients with other clinical characteristics. Therefore in the following only estimates of probabilities will be used which have been based on a regression model which accounts for the dependence between the diagnostic variables.

11.3 Referral for non-invasive tests prior to arteriography

The protocol has to advise, based on clinical data, if non-invasive testing is indicated, and if affirmative, which sequence is the best, in order to select the patients with the highest chance of having coronary artery disease. Thus, to develop a diagnostic protocol for non-invasive testing it is necessary to get an answer to the following questions:

- Which, if any, non-invasive test should be performed for optimal referral for arteriography?,
- If combinations of tests have to be performed what is the optimal sequence?,
- Will a non-invasive diagnostic step be performed irrespective of its costs or will there be a threshold for the cost-effectiveness ratio?

11.4 Criterion for non-invasive tests

There are categories of patients for whom non-invasive testing does not alter the decision to perform an arteriogram and for whom non-invasive testing is not useful. If the $P_{threshold}$ is determined, these categories of patients can be identified, namely those for whom the pre-test probability and post-test probability, after all possible non-invasive tests, are both higher or both lower than the established $P_{threshold}$.

For example, males older than 60 years with typical angina pectoris have, based on our data, a pre-test probability of 92% (Figure 11.3). If all three non-invasive tests are negative, the post-test probability is still 56%, which is $>P_{threshold}$ (=15%). Thus, all these patients are still candidates for arteriography. Non-invasive testing does not alter that decision and therefore non-invasive testing is of no value in this case.

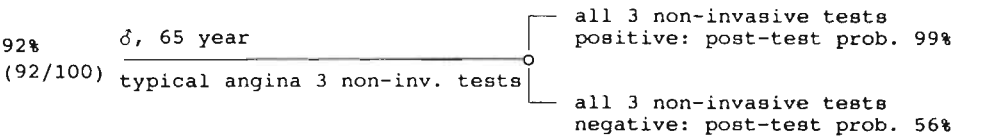


Figure 11.3 Pre- and post-test probabilities for males with typical angina, > 60 years old with all 3 non-invasive tests either positive or negative.

It is more complicated for those patients for whom, depending on the test result, the decision of referral for arteriography changes (Figure 11.4). In general, no arteriography should be performed in females < 50 years with atypical angina as the pre-test probability, based on our data, is only 6%. However, if fluoroscopy is

performed, a number of patients may have a positive test and for them the post-test probability becomes 33%. These latter patients are now candidates for arteriography and the non-invasive testing has changed the decision. The same is true for males 50-60 years with nonspecific chest pain/no symptoms and a negative fluoroscopy test. The post-test probability is reduced to <15% (Figure 11.4). These patients would now be denied arteriography.

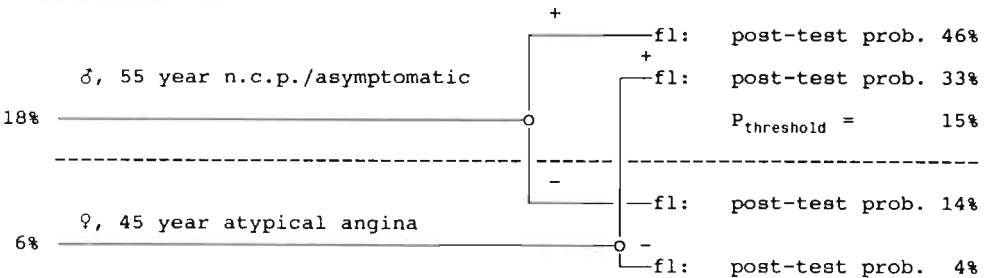


Figure 11.4 Pre- and post-test probabilities for 2 categories of patients with (cine) fluoroscopy applied as non-invasive test; $P_{\text{threshold}}$ set at 15%

The same applies if only the combination of the results of two tests changes the referral decision. For example, in females younger than 50 years with typical angina, only if fluoroscopy and stress ECG are both performed and are both negative, is arteriography not advised (Chapter 11, Appendix 3, Figure A3). Whether it is rational to perform a test as a step in the diagnostic process, it is necessary to know the probabilities before and after this step. Also these probabilities can be estimated by logistic regression by excluding those variables from the model that are related to tests still not performed. In this way it was determined for all categories which tests or test combination proved to be rational and this is listed in Table 11.2.

Table 11.2 Tests or test combinations which proved to be rational for various categories of patients

Category	
Female, <50 years	
n.c.p./asymptomatic	Fluoroscopy combined with stress ECG
atypical angina	Fluoroscopy combined with stress ECG
typical angina	Fluoroscopy combined with stress ECG
Female, 50-59 years	
n.c.p./asymptomatic	Fluoroscopy
atypical angina	Fluoroscopy combined with stress ECG
typical angina	Fluoroscopy
Female, ≥ 60 years	
n.c.p./asymptomatic	Fluoroscopy
atypical angina	Fluoroscopy combined with stress ECG
typical angina	Direct referral for arteriography
Male, < 50 years	
n.c.p./asymptomatic	Fluoroscopy combined with stress ECG
atypical angina	Fluoroscopy combined with stress ECG and thallium scintigraphy
typical angina	Direct referral for arteriography
Male, 50-59 years	
n.c.p./asymptomatic	Fluoroscopy combined with stress ECG
atypical angina	Direct referral for arteriography
typical angina	Direct referral for arteriography
Male, ≥ 60 years	
n.c.p./asymptomatic	Fluoroscopy combined with stress ECG
atypical angina	Direct referral for arteriography
typical angina	Direct referral for arteriography

11.5 Determination of the optimal sequence of tests

When for a given category of patients a certain combination of non-invasive tests appears useful, one has to decide on the optimal sequence. After performing fluoroscopy and thallium scintigraphy, independently of the sequence of these tests, the same patients will be referred for arteriography, namely those who have a post-test probability based on the two test results higher than 15% (Figure 11.5).

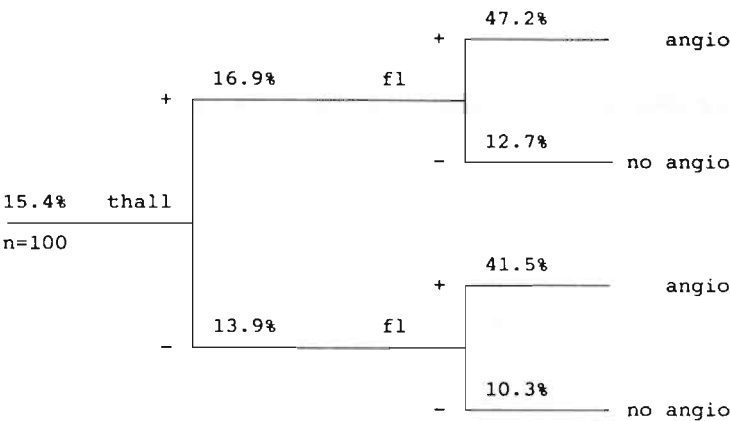


Figure 11.5.a

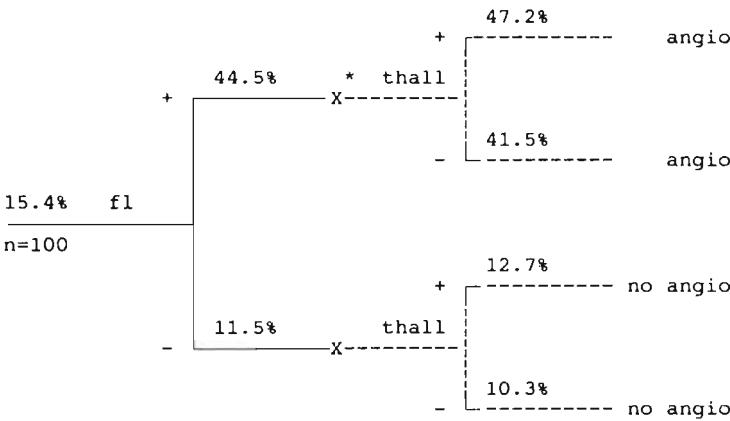


Figure 11.5.b.

Figure 11.5.a and b Two different sequences for fluoroscopy and thallium scintigraphy in females, 50 - 59 years old and with atypical angina

* The dotted lines in the flow charts indicate diagnostic tests of which the results do not influence the decision for referral for arteriography and are consequently not advised

Females who are 50-59 years old and have atypical angina have a pre-test probability of 15.4% and would have been referred for arteriography if no further testing were done. As shown in Figure 11.5.a, after performing a thallium scintigraphy, independent of the result of this test, there is still a reason for

performing a fluoroscopy. Thus both tests have to be performed on all 100 patients. However, if fluoroscopy is performed first, the result of the thallium scintigraphy no longer influences the decision whether or not to refer for arteriography (Figure 11.5.b). In this latter sequence it is only necessary to perform fluoroscopy on the 100 patients. With both sequences of tests the same patients will be referred for arteriography but in the second sequence there is no need for a thallium scintigraphy with its inconveniences and hazards for the patients and its costs. This conclusion was already derived from Table 11.2 without considerations of the costs. However, it can be less apparent as the next case will show. In Figures 11.6.a & b, both sequences for fluoroscopy and stress ECG are shown for the same category of patients. If the result of the test no longer influences the referral decision, it is not included. In both sequences the same patients are refused an arteriogram, namely those with a negative fluoroscopy test, as well as a non-diagnostic or negative stress ECG. If one starts with fluoroscopy, this test has to be done in all 100 patients and stress ECG 88 (=88.2) times. In the reverse sequence, stress ECG has to be performed in all 100 patients and fluoroscopy only 83 (=41.5+41.4) times.

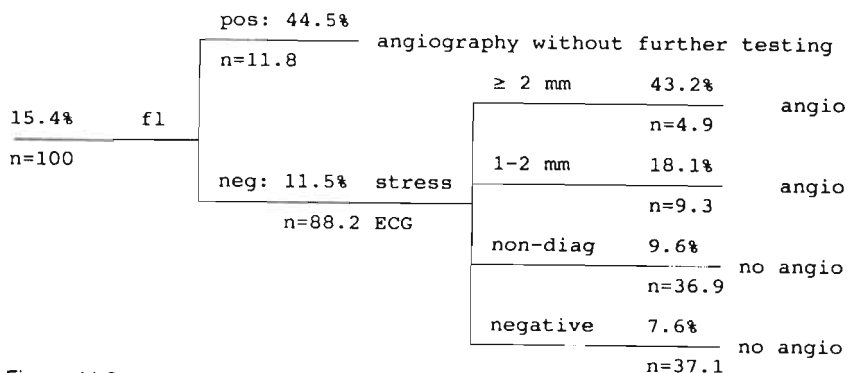


Figure 11.6.a

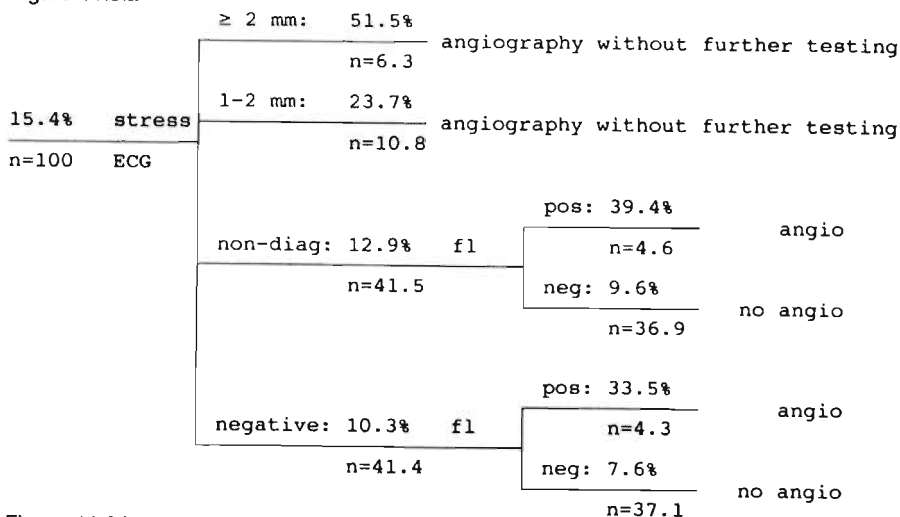


Figure 11.6.b

Figure 11.6.a and b Two different sequences for fluoroscopy and stress ECG in females, 50 -59 years old and with atypical angina

Without knowledge of the inconveniences, hazards and costs of both tests, it is not possible to decide which sequence has the preference. In the following only the costs of the test are taken into account. As discussed in Chapter 11, Appendix 1 the cost of performing fluoroscopy once is set at Dfl. 12.5 and the cost of a stress ECG is set at Dfl. 100. The first sequence applied to the 100 patients costs:

$$100 \times \text{Dfl. 12.5} + 88.2 \times \text{Dfl. 100} = \text{Dfl. 10,700}$$

and the second sequence costs:

$$100 \times \text{Dfl. 100} + 82.9 \times \text{Dfl. 12.5} = \text{Dfl. 11,036}$$

Although in this case the difference in costs is minimal the first sequence is more advantageous and is on the average per patient about 40 guilders cheaper. This approach can also be applied to sequences of more than two tests, and by including the costs of the tests the most cost effective sequence can be determined. Furthermore as can be seen from Figure 11.6 it is necessary to be informed about the number of times each test result occurs. The estimates for these numbers are provided by the log-linear model as discussed in the appendix of Chapter 5.

11.6 Determination of the cost-effectiveness ratio of a non-invasive diagnostic step

The last question which has to be answered is whether the ratio between costs and effectiveness of a diagnostic step could influence the choice whether or not to perform the test. Apart from the costs it is also necessary to have a kind of measure of the effectiveness of a diagnostic step.

There are two characteristics of a diagnostic step that influence its effectiveness:

- (1) The post-test probability value: the nearer to 100% or 0%, the more justification to refer or not to refer the patient for arteriography.
- (2) The percentage of patients in whom the non-invasive diagnostic step changes the referral decision. If for instance the post-test probability after a positive result is 99.9%, performance of this step may still be reconsidered if this situation occurs only once in 10,000 patients.

A method for producing an expression for the effectiveness of a test that accounts for these two factors is described in the Appendix 2 of this chapter. Here this method will be demonstrated in the two following examples. In the first example the effectiveness of (cine) fluoroscopy is determined for females, ≥ 60 years, with atypical angina (Figure 11.7).

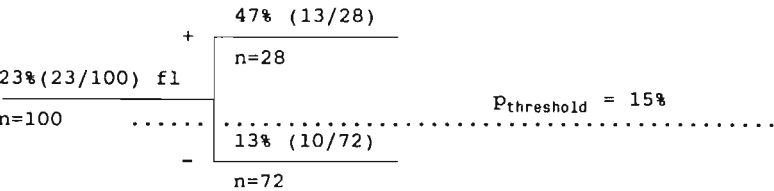


Figure 11.7 Flow chart for female, ≥ 60 years, atypical angina and (cine)fluoroscopy, with the number of patients and the probability in percentage
23% = pre-test probability based on sex, age and symptoms, while accounting for dependence between these clinical variables

The pre-test probability of coronary artery disease is 23%. With a $P_{\text{threshold}}$ of 15%, all patients in this category would have arteriography, if one refrains from further non-invasive testing. If arteriography is performed in 100 patients, the arteriogram will be positive in 23 patients and these patients will have access to therapy. Due to this therapy the clinical status of 23 patients with coronary artery disease would in principle improve. As shown in Chapter 5, the benefit, called B_{therapy} , is defined by the improvement of the quality of life or increase in life expectancy thanks to this therapy, less the risks and costs of this therapy and those of the arteriography. This benefit is expressed in B_{therapy} units and in this example applying arteriography in all 100 patients will result in a benefit due to therapy of 23 units. However, arteriography was performed unnecessarily in 77 patients. Performing an arteriogram on a healthy person has certain negative effects that should be considered such as the risk of mortality or the morbidity due to the arteriogram, the inconveniences and the cost of the arteriogram. This loss is called $L_{\text{arteriography}}$ and is expressed in $L_{\text{arteriography}}$ units. In this example 77 $L_{\text{arteriography}}$ units are on the debit side and 23 B_{therapy} units will be on the credit side. To determine the net benefit of the arteriography in these 100 patients, the debit must be subtracted from the credit, and it is thus necessary to transform $L_{\text{arteriography}}$ units into B_{therapy} units. If the post-test probability is equal to $P_{\text{threshold}}$ ($= 15\%$), then the decision whether or not to perform an arteriography is of no importance whatsoever, suggesting that the benefit of the correct referrals for arteriography is equal to the loss due to the incorrect ones. In 100 patients with a probability of coronary artery disease of 15% ($=P_{\text{threshold}}$) 15 referrals would be correct and 85 referrals would be incorrect. Therefore,

$$\begin{aligned}
 P_{\text{threshold}} \times B_{\text{therapy}} &= (1-P_{\text{threshold}}) \times L_{\text{arteriography}} \quad (\text{Appendix 11.2}) \\
 \Rightarrow 15 \times B_{\text{therapy}} &= 85 \times L_{\text{arteriography}} \\
 \Rightarrow L_{\text{arteriography}} &= 15/85 B_{\text{therapy}}
 \end{aligned}$$

Thus, one $L_{\text{arteriography}}$ unit equals $0.176 B_{\text{therapy}}$ unit.

Returning to our example 77 referred patients did not have coronary artery disease and 77 $L_{\text{arteriography}}$ units equal $13.6 B_{\text{therapy}}$ units ($77 \times (15/85)$). 23 patients have after the correct referral the benefit of therapy and the net benefit of arteriography for these 100 patients is 9.4 ($23-13.6$), expressed in B_{therapy} units. Returning to the primary question, What is the value of fluoroscopy as a non-invasive test for these patients? This test will be positive in 28 patients, with a post-test probability of 47% and negative with a post-test probability of 13% for 72 patients (Figure 11.7) Arteriography will now be performed only in these 28 patients and disease will be present in 13 of them (0.47×28). Thus, on the credit side we will obtain 13 B_{therapy} units, and on the debit side, 15 $L_{\text{arteriography}}$ units, which equals $2.6 (15 \times 0.176) B_{\text{therapy}}$ units. 72 patients will be denied arteriography, and for them no benefits or losses due to arteriography have to be accounted for, i.e this does not apply to them. Thus, after fluoroscopy the net benefit obtained by arteriography is $10.4 B_{\text{therapy}}$ units ($13-2.6$).

The additional value of fluoroscopy over non-testing for these 100 patients is the net benefit of arteriography after performing fluoroscopy minus the net benefit of

arteriography without fluoroscopy. In the appendix this additional value or effectiveness of performing a non-invasive test once is called E_{test} . In this example the E_{test} expressed in B_{therapy} units would be the difference of 10.4 and 9.4, divided by 100, equals 0.01 unit.

The situation is different in the second example for females, < 50 years, with atypical angina (Figure 11.8) where the pre-test probability of coronary artery disease (9%) is lower than $P_{\text{threshold}}$.

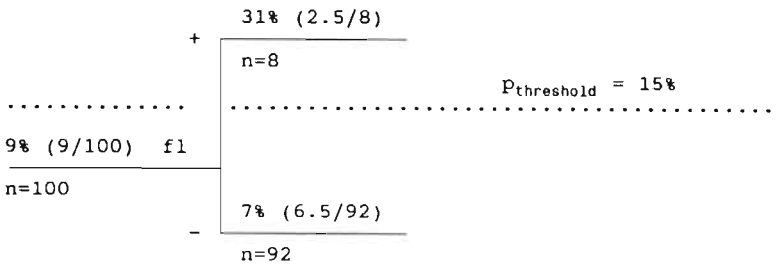


Figure 11.8 Flow chart for female, < 50 years, atypical angina and (cine) fluoroscopy, with the number of patients and the probability as a percentage
9% = pre-test probability based on sex, age and symptoms, while accounting for dependence between these clinical variables

This implies that, without non-invasive testing arteriography would be withheld from all these patients. However, if fluoroscopy is performed this test is positive in 8 patients with a post-test probability of 31%. Arteriography would be advised for these 8 patients and coronary artery disease would be present in 2.5 individuals (0.31×8). Due to the performance of fluoroscopy, 2.5% of the patients now have access to therapy. Thus, the value of fluoroscopy testing is 2.5 B_{therapy} units minus 5.5 $L_{\text{arteriography}}$ units and using the same converting factor this value is equal to $2.5 - 5.5 \times 0.176 = 1.5 B_{\text{therapy}}$ units. Without fluoroscopy no patients would have been referred for arteriography so there would be no benefits or losses. For these 100 patients fluoroscopy results in a net profit of 1.5 B_{therapy} units and so the net result is 1.5 minus zero = 1.5 B_{therapy} unit. The E_{test} of fluoroscopy for these patients is $1.5/100 = 0.015 B_{\text{therapy}}$ units. For the situation with a pre-test probability higher than the $P_{\text{threshold}}$, a formula for the value of a test, E_{test} , is derived in the appendix 11.2. With P and N designating the pre-test probability and number of patients, and with P_2 and N_2 the post-test probability and the number of the patients with a negative test result, this formula is

$$E_{\text{test}} = N_2/N \times (P_{\text{threshold}} - P_2)/(1 - P_{\text{threshold}}) \quad (11.1)$$

This expression clarifies that the effectiveness of a test is determined by two factors as stated at the beginning of this section. One is the fraction for whom the referral is withheld as a consequence of the test result. This is expressed by the factor N_2/N . The value of a test is also determined by the level of certainty with which the decision to defer is taken. This is expressed by the second part of the expression and it can be seen that this depends on the difference between the post-test probability P_2 and the $P_{\text{threshold}}$. For the situation with a pre-test probability lower than the $P_{\text{threshold}}$, in Appendix 11.2 a similar formula has been

derived. With P_1 and N_1 representing the probability and the number of patients with a positive test result, this formula is

$$E_{\text{test}} = N_1/N \times (P_1 - \text{Odds}_{\text{threshold}})/(1 - P_{\text{threshold}})$$

and the same considerations apply.

In conclusion it can be seen that to determine this E_{test} it is necessary to know the pre- and post-test probability estimates as well as the proportion of the patients with a positive and negative result of the non-invasive test.

In section 11.5 (and in particular in appendix 1 of this chapter) it has been shown how the costs of a non-invasive test can be assessed. By relating these costs to the effectiveness of a test, a cost-effectiveness ratio can be calculated.

With help of formula 11.1, it can be calculated for the example of Figure 11.6.a that E_{test} equals

$$\begin{aligned} &= 88.2/100 \times (0.15 - 0.115)/(1 - 0.15) \\ &= 0.036 B_{\text{therapy}} \text{ unit} \end{aligned}$$

The cost of fluoroscopy is Dfl. 12.5 and so the cost-effectiveness ratio of a fluoroscopy test for females, 60 years and older and atypical angina is about $12.5/0.036 = 344$ guilders per B_{therapy} unit. The same calculation can be made for the sequence of fluoroscopy and stress ECG as one diagnostic procedure.

The E_{test} of the sequence is determined as follows:

$$\begin{aligned} &= 36.9/100 \times (0.15 - 0.096)/(1 - 0.15) \\ &+ 37.1/100 \times (0.15 - 0.076)/(1 - 0.15) \\ &= 0.056 B_{\text{therapy}} \text{ unit} \end{aligned}$$

As shown in appendix 1 of this chapter, the cost of fluoroscopy and stress ECG is Dfl. 112.50 and thus the cost effectiveness ratio equals 2,008 guilders per B_{therapy} unit.

In this way the cost-effectiveness ratio for each non-invasive diagnostic step or sequence of steps can be calculated and a comparison of different diagnostic steps or sequences of steps is now possible in a more quantitative way.

With the help of a personal computer and specially developed software, the cost-effectiveness ratios have been calculated for all possible combinations and sequences of the three non-invasive tests in all categories of patients. Consecutively the optimal sequence for each category has been determined. The sequence of tests with the best cost-effectiveness ratio proved to be fluoroscopy as first non-invasive test, followed by stress ECG testing, and thallium scintigraphy as third.

The value of the cost-effectiveness ratio of fluoroscopy in females aged 50-59 years with atypical angina, appeared to be the lowest ratio for all categories of patients and all possible non-invasive tests and was Dfl. 344 per B_{therapy} unit. In contrast, thallium scintigraphy proved to change the management only in one single category of patients namely, males aged < 50 years with atypical angina (Figure 11.2) and the cost-effectiveness ratio of thallium scintigraphy in this situation proved to be more than Dfl. 150,000 per B_{therapy} unit.

For females less than 50 years old with typical angina, only the combination of fluoroscopy and stress ECG could influence the arteriographic referral decision

and the cost-effectiveness ratio of these both tests combined proved to be Dfl. 30,000 per B_{therapy} unit.

These diagnostic steps had by far the worst cost-effectiveness ratio. As explained in this subsection, the reason for these relative high values can be linked to post-test probabilities that are just barely under the margin of 15%.

Compared with the lowest ratio of Dfl. 344 per unit, cost-effectiveness ratios of Dfl. 150,000 and Dfl. 30,000 per B_{therapy} unit are considered excessively high. For this reason neither diagnostic step was included in the proposal for the diagnostic protocol. As a consequence, any use of thallium scintigraphy was ruled out in the diagnostic work-up of patients with symptoms suggesting coronary artery disease. One has to remember, however, that this conclusion is based on data on thallium scintigraphy of which the validity has been questioned (as discussed in Chapter 8).

The resulting protocols of non-invasive testing for all different categories of patients are recorded in the appendix of this chapter (Appendices 3 and 4) and are schematically shown in Figure 11.9.

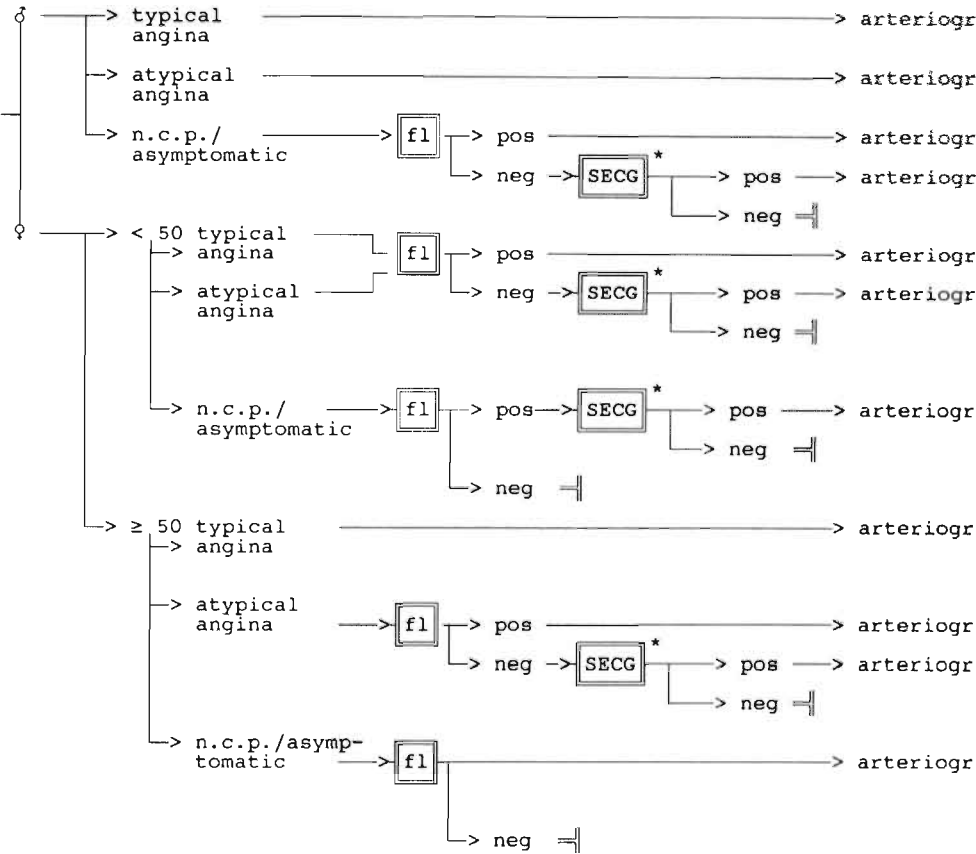


Figure 11.9 Protocol as suggested for the different categories of patients, regarding the use of fluoroscopy and stress ECG testing, and taking dependence into account.

pos = positive test result / neg = negative test result; fl = Fluoroscopy; SECG = Stress ECG; * since stress ECG does have more than two test results, the terms 'neg' and 'pos' should be interpreted according to the flow charts, given for each category in Chapter 11, appendix 3; = = end of protocol

Chapter 11, Appendix 1

Estimation of the costs of the different non-invasive tests and cine coronary arteriography

The real costs of the investigations are not known. These would involve costs of personnel, material and equipment with its various aspects as purchase, maintenance, etc. Relevant information is still lacking at this phase of reorganisation of the financial structure of the health care in the Netherlands.

Prior to the present system of fixed budgets the costs were charged to insurance companies on a basis of a fee for service.

These charges consisted of a physician fee and costs refunded to the hospital. The following figures are based on these charges and have been used in our analysis.

Cost for the various investigations as charged by the hospital:

Stress ECG	: Dfl. 15
Thallium scintigraphy	: Dfl. 700
Coronary arteriogram (costs of hospitalisation included)	: Dfl. 2000
Fluoroscopy to detect calcification in coronary arteries:	00
(no costs chargeable for fluoroscopy, as costs are already included in the costs of the chest X-ray which is always indicated for these patients)	

The physician's fees for the various investigations have also to be included in the cost. The cost of the investigations are then as follows:

Stress ECG	: Dfl. 100
Thallium scintigraphy	: Dfl. 850
Coronary arteriogram	: Dfl. 2800
Fluoroscopy	: Dfl. 12.5

These financial costs may vary from hospital to hospital in the Netherlands and these figures only apply to the De Wever Hospital.

Chapter 11, Appendix 2

The averaged benefit of performing one non-invasive test expressed in $B_{therapy}$ units

As deduced from the diagram in Chapter 5 the following equation (5.4) for $P_{threshold}$ holds:

$$P_{threshold} \times N \times B_{therapy} = (1 - P_{threshold}) \times N \times L_{arteriography}$$

or

$$L_{arteriography} = B_{therapy} \times P_{threshold} / (1 - P_{threshold}) \quad (A11.1)$$

Furthermore, in Chapter 5 an expression (5.3) for the benefit of a referral for arteriography (and its consequences) over non-referral for N patients was derived:

$$\text{Benefit} = P \times N \times B_{therapy} - (1 - P) \times N \times L_{arteriography} \quad (A11.2)$$

By substituting $L_{arteriography}$ from A11.1 in equation A11.2 the benefit can now be expressed in $B_{therapy}$ units:

$$\begin{aligned} \text{Benefit} &= P \times N \times B_{therapy} - (1 - P) \times N \times B_{therapy} \times P_{threshold} / (1 - P_{threshold}) \\ &= N \times B_{therapy} \times (P - (1 - P) \times P_{threshold} / (1 - P_{threshold})) \\ &= N \times B_{therapy} \times (P - P_{threshold}) / (1 - P_{threshold}) \end{aligned} \quad (A11.3)$$

As this is the net benefit for N patients, the average benefit for one patient can be determined by dividing A11.3 by N. This is the averaged benefit of one referral for arteriography and equals:

$$B_{therapy} \times (P - P_{threshold}) / (1 - P_{threshold}) \quad (A11.4)$$

As the benefit will be expressed in $B_{therapy}$ units A11.4 reduces to

$$(P - P_{threshold}) / (1 - P_{threshold}) \quad (A11.5)$$

To determine the averaged benefit of performing one non-invasive test two situations have to be distinguished: the situation where the pre-test probability is higher and the situation that the pre-test probability is lower than the $P_{threshold}$.

In the situation that the pre-test probability is higher than the $P_{threshold}$ (Figure A11.1) all patients would be candidates for arteriography.

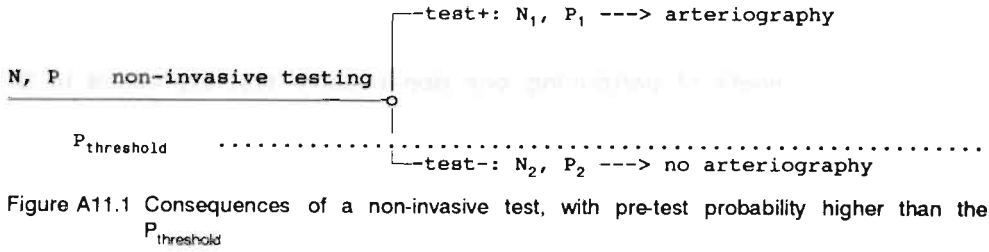


Figure A11.1 Consequences of a non-invasive test, with pre-test probability higher than the $P_{\text{threshold}}$

In this figure:

N = number of patients in whom the test is performed; P = probability of disease before testing; P_1 = probability of disease after a positive test result; N_1 = number of patients with a positive test result; P_2 = probability of disease after a negative test result; N_2 = number of patients with a negative test result

If no non-invasive test is performed, the benefit of referral for arteriography for these N patients expressed in B_{therapy} units, will be:

$$N \times (P - P_{\text{threshold}}) / (1 - P_{\text{threshold}}) \quad (\text{A11.6})$$

However, healthy patients are included in this group. To exclude these patients as much as possible, a non-invasive test is performed. Depending on the test result the post-test probability for a portion of patients N_1 rises to P_1 and for the complementary portion of patients N_2 the probability P_2 falls below the $P_{\text{threshold}}$. The benefit of the N_1 patients with a positive test result and a probability P_1 can also be calculated with formula A11.3 and equals

$$N_1 \times (P_1 - P_{\text{threshold}}) / (1 - P_{\text{threshold}}) \quad (\text{A11.7})$$

Testing only has consequences for the patients with a negative test result as these patients will not be referred for arteriography. Consequently the term for the benefit for these patients is irrelevant and is omitted.

The benefit as a consequence of (non-invasive) testing compared with the benefit if the test is not performed forms the value of the test and will be expressed as E_{test} . If the test is performed N times this value equals $N \times E_{\text{test}}$ and can be calculated by subtracting expression A11.6 from expression A11.7

$$N \times E_{\text{test}} = N_1 \times (P_1 - P_{\text{threshold}}) / (1 - P_{\text{threshold}}) - N \times (P - P_{\text{threshold}}) / (1 - P_{\text{threshold}}) \quad (\text{A11.8})$$

By performing the non-invasive test the total number of patients does not change, thus,

$$N = N_1 + N_2 \quad (\text{A11.9})$$

Furthermore, the total number of patients with disease equals $P \times N$ and this number also does not change and therefore

$$P \times N = P_1 \times N_1 + P_2 \times N_2 \quad (\text{A11.10})$$

Substituting A11.9 and A11.10 in A11.8 and dividing the result by N gives

$$E_{\text{test}} = N_2/N \times (P_2 - P_{\text{threshold}})/(1-P_{\text{threshold}}) \quad (\text{A11.11})$$

and this variable is also expressed in B_{therapy} units.

In the situation that the pre-test probability is lower than the $P_{\text{threshold}}$ (Figure A11.2) no patients would be candidates for arteriography

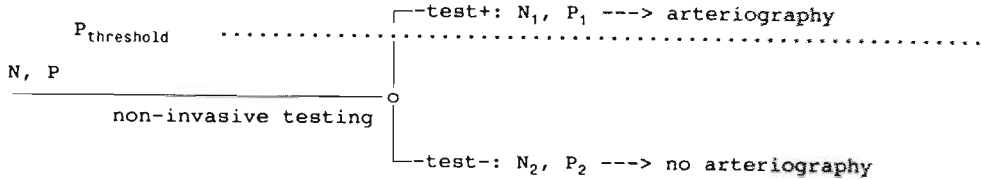


Figure A11.2 Consequences of a non-invasive test, with pre-test probability lower than the $P_{\text{threshold}}$. (for the explanation of the symbols see Figure A11.1)

The probability of having coronary artery disease without testing is now below $P_{\text{threshold}}$, thus none of the patients will have an arteriogram and consequently none of the patients will profit from invasive therapy (B_{therapy}).

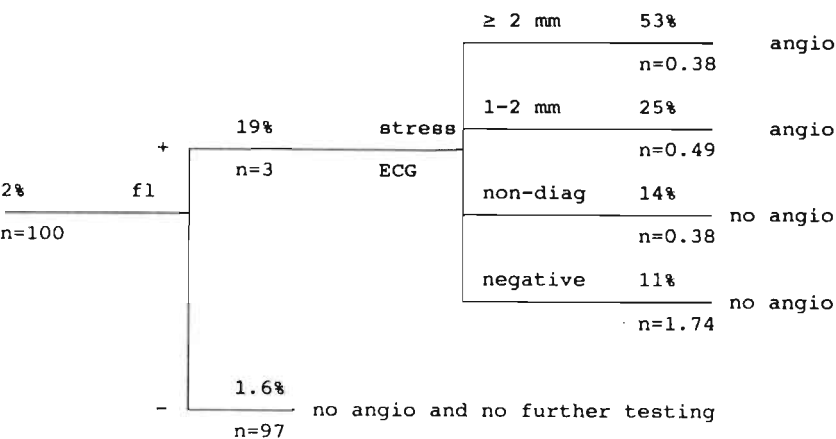
After testing, a fraction of patients (N_1) have a post test-probability (P_1) higher than the $P_{\text{threshold}}$, and will now undergo an arteriogram; some of these patients will prove to have coronary artery disease, others not.

The profit of this testing can be deduced in the same way as in the situation where the pre-test probability is higher than the $P_{\text{threshold}}$:

$$E_{\text{test}} = N_1/N \times (P_1 - P_{\text{threshold}})/(1-P_{\text{threshold}}) \quad (\text{A11.12})$$

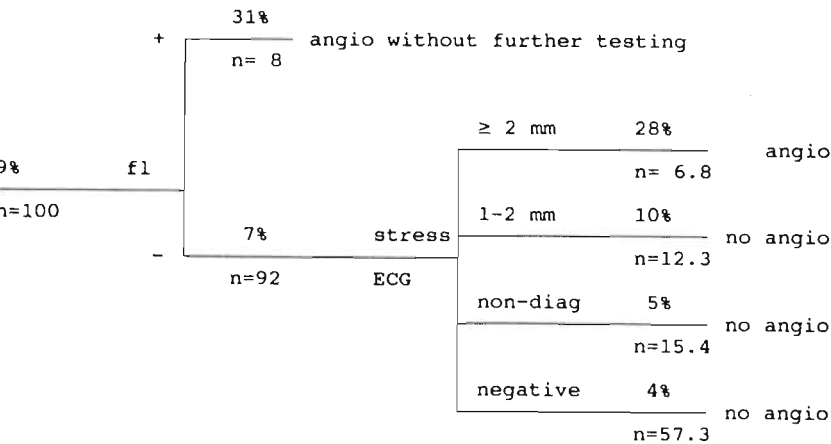
Chapter 11, Appendix 3

Flow charts

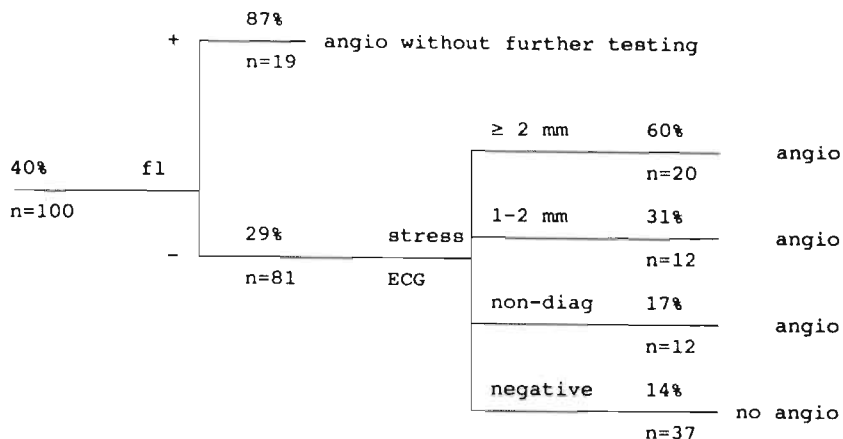


(A1) Female, < 50 years, n.c.p./asymptomatic, with the number of patients and the probability in percentage

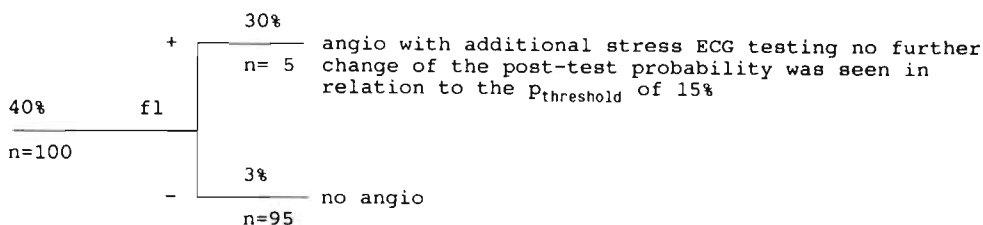
2% = pre-test probability of having coronary artery disease, based on clinical variables, while taking dependence into account; n=100 = number of (100) patients; 19% = post-test probability, with a positive fluoroscopy test (fl); n=4 = number of patients (4) with a positive fluoroscopy test; 53% = post-test probability with a ST segment depression of ≥ 2 mm; n=0.38 = number of patients with this test result per 100 patients



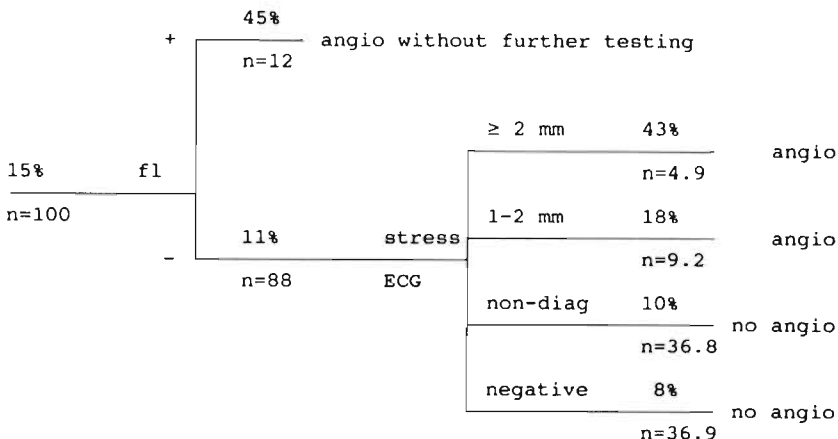
(A2) Female, < 50 years, atypical angina.



(A3) Female, < 50 years, typical angina

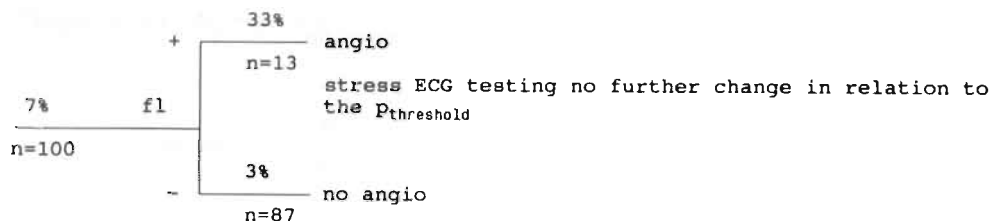


(A4) Female, 50 - 59 years, n.c.p./asymptomatic

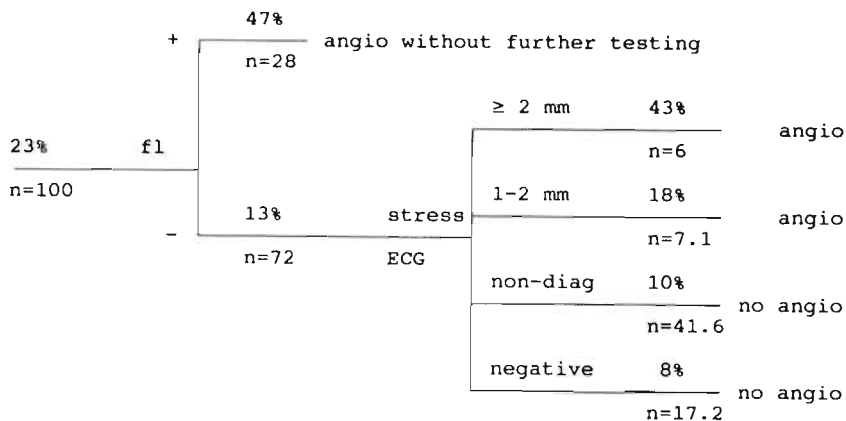


(A5) Female, 50 - 59 years, atypical angina

(A6) Female, 50 - 59 years, typical angina. The pre-test probability for this category is 55%. Result of either non-invasive test fails to reduce the post-test probability below 15%.



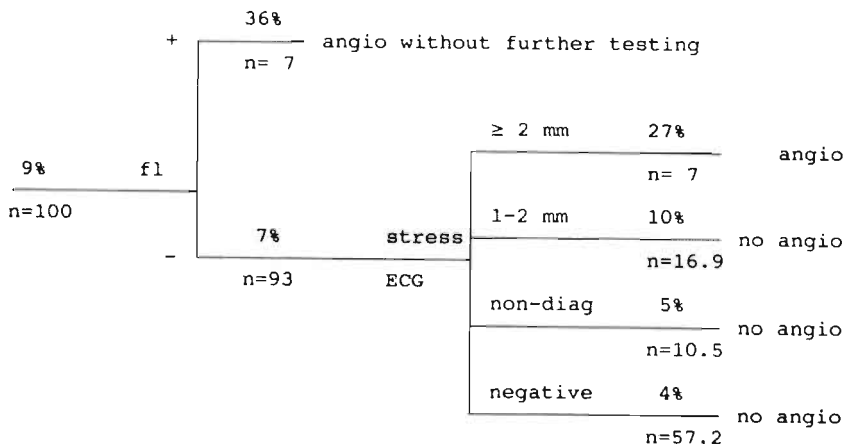
(A7) Female, ≥ 60 years, n.c.p./asymptomatic



(A8) Female, ≥ 60 years, atypical angina

The pre-test probability for this category is 71%. Result of either non-invasive test fails to reduce the post-test probability below 15%.

(A9) Female, ≥ 60 years, typical angina



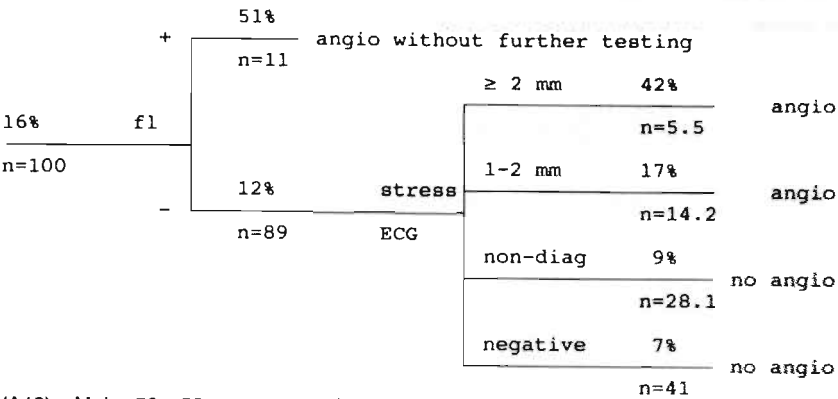
(A10) Male, < 50 years, n.c.p./asymptomatic

The pre-test probability for this category is 29%. Result of either non-invasive test fails to reduce the post-test probability below 15%.

(A11) Male, < 50 years, atypical angina

The pre-test probability for this category is 73%. Result of either non-invasive test fails to reduce the post-test probability below 15%.

(A12) Male, < 50 years, typical angina



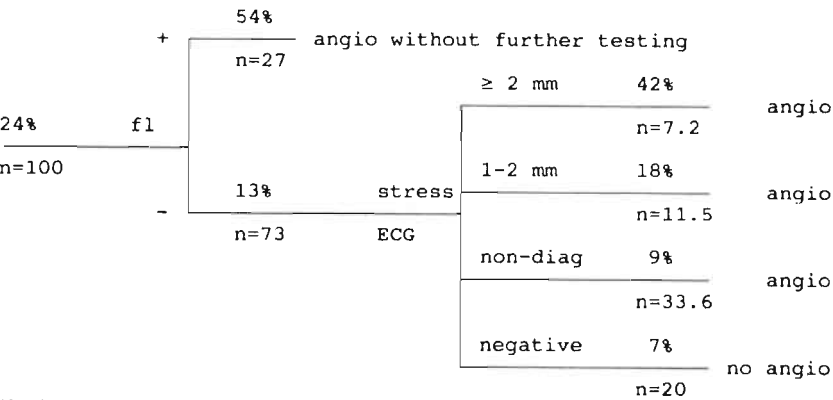
(A13) Male, 50 - 59 years, n.c.p./asymptomatic

The pre-test probability for this category is 44%. Result of either non-invasive test fails to reduce the post-test probability below 15%.

(A14) Male, 50 - 59 years, atypical angina

The pre-test probability for this category is 84%. Result of either non-invasive test fails to reduce the post-test probability below 15%.

(A15) Male, 50 - 59 years, typical angina



(A16) Male, ≥ 60 years, n.c.p./asymptomatic

The pre-test probability for this category is 54%. Result of either non-invasive test fails to reduce the post-test probability below 15%.

(A17) Male, ≥ 60 years, atypical angina

The pre-test probability for this category is 91%. Result of either non-invasive test fails to reduce the post-test probability below 15%.

(A18) Male, ≥ 60 years, typical angina

Chapter 11, Appendix 4

The use of the following protocol as suggested for the different categories (shown schematically in Figure 11.9):

Female, < 50 years	
n.c.p./asymptomatic	: fluoroscopy, if positive followed by stress ECG testing
atypical angina	: fluoroscopy, if negative followed by stress ECG testing
typical angina	: arteriography only
Female, 50 - 59 years	
n.c.p./asymptomatic	: fluoroscopy
atypical angina	: fluoroscopy, if negative followed by stress ECG testing
typical angina	: arteriography only
Female, ≥ 60 years	
n.c.p./asymptomatic	: fluoroscopy
atypical angina	: fluoroscopy, if negative followed by stress ECG testing
typical angina	: arteriography only
Male, < 50 years	
n.c.p./asymptomatic	: fluoroscopy, if negative followed by stress ECG testing
atypical angina	: arteriography only
typical angina	: arteriography only
Male, 50 - 59 years	
n.c.p./asymptomatic	: fluoroscopy, if negative followed by stress ECG testing
atypical angina	: arteriography only
typical angina	: arteriography only
Male, ≥ 60 years	
n.c.p./asymptomatic	: fluoroscopy, if negative followed by stress ECG testing
atypical angina	: arteriography only
typical angina	: arteriography only

12 | Evaluating a protocol in practice

12.1 Introduction

It seems only right that having constructed a protocol, it should be evaluated in practice. This should be done by comparing the patient management resulting from the application of the protocol with patient management as practised by cardiologists in the De Wever Hospital. Furthermore, for a proper comparison this should be done prospectively with a new patient population.

The essential element of this comparison will be the possible differences in the referral pattern for arteriography and how these correlate with disease status.

Furthermore, the protocol recommends a specific work-up with the help of non-invasive testing for each category of patients. As a possible secondary effect any reduction in the number of non-invasive tests could also be evaluated.

The protocol is based on the estimates of disease probabilities of the log-linear model. The coefficients of this model have been calculated with the help of data collected over the years 1983-1987. However, it is possible that the correlation between diagnostic characteristics and disease status has drifted in the course of time and this may impair the correctness of the probability estimates as provided by the log-linear model. By comparing the probabilities as expected by the log-linear model with the probabilities as observed in this prospective study, it is possible to check for such a possible change. Although indirectly, this is also an evaluation of the validity of the protocol.

A final but no less important issue is the detection of calcification in the coronary arteries by fluoroscopy, which plays an important role in the protocol. However, the cine coronary arteriogram report, thus cine fluoroscopy, was the basis of our data collection for the detection of calcification. This method differs from "simple" fluoroscopy, the usual way of performing this non-invasive examination in clinical practice. Thus, it is essential to know if the discriminating value of simple fluoroscopy differs from that of cine fluoroscopy. This will be assessed additionally in the prospective study.

Design of the study population for the evaluation of the protocol

The first evaluation method to consider would be an experiment with a non-selective partitioning of all the patients referred for arteriography into two groups. The work-up of one group would be performed as advised by the protocol, and of the other group by the existing procedure. The results of both procedures could then be compared, using a proper outcome criterion. This method, however, is not very practical for a number of reasons. It would imply that according to the

protocol a patient is denied an arteriogram which has been judged as indicated by the cardiologist. Or, vice versa, the protocol advises the performance of an arteriogram in contrast to the opinion of the cardiologist. As the cardiologist is ultimately responsible for the management of the patient, his permission is needed for an alternative protocol. When both groups of patients are handled at the same hospital the future referral pattern of the cardiologists may be influenced due to the knowledge of the protocol and this may well diminish its effect. Therefore, another and in fact more simple approach was chosen.

The population will only consist of patients referred for arteriography for diagnostic reasons as judged by the cardiologists. Thus, applying the protocol does not add extra referrals for arteriography. This definition of the study population implies that the protocol is validated in only a subset of the population of patients consulting the cardiologist for symptoms suggesting ischaemic heart disease. This is a minor predicament because, as discussed in Chapter 2, it seems reasonable to assume that patients who are not referred for arteriography by the cardiologist are in no need of such an investigation. Furthermore, the cardiologist will perform all non-invasive tests he considers necessary. Dependent on the category of patients, it may occur that the protocol recommends a stress ECG and/or fluoroscopy and that these examinations have not yet been requested by the referring cardiologist.

As a stress ECG can only be performed after authorization by a cardiologist, permission to perform one is necessary for each patient. This also could give the cardiologist a clue to the work-up as advised by the protocol, possibly influencing his referral pattern for future patients. Therefore, in these patients additional stress testing will not be performed. This problem does not exist for additional fluoroscopy as this is performed by a radiologist. A general permission from the cardiologists to do so was requested and obtained but the additional investigation was performed only after informed consent from the patient. Permission was granted by the Medical Ethics Committee of the De Wever Hospital for this procedure.

Execution of the study

The patients taking part in the prospective study are those who have been referred by one of the practising cardiologists for a diagnostic cine coronary arteriogram, that is to say an arteriogram to exclude or to confirm coronary artery disease. Thus, patients with a proven myocardial infarction are excluded. The same applies for patients with a previous abnormal arteriogram. Patients with valvular or congenital heart disease are also exempted, and those with unstable angina are also excluded from the study for the same reason as discussed in Chapter 2.

The cardiologists refer the patients for arteriography and for eventual prior non-invasive testing, without any knowledge of the protocol. If coronary disease is proven to be present, a stress ECG may still be indicated in some patients to evaluate the exercise tolerance of the patient. Thus, not all stress ECGs indicated as "unnecessary" according to the protocol can be labelled as such and it would be wrong to conclude that costs of all those "unnecessary" stress ECGs could have been avoided. Therefore, the referring cardiologist is asked to indicate if the stress ECG, if performed, was carried out to confirm or exclude coronary artery disease (a diagnostic investigation) or as an evaluation of the exercise tolerance.

In summary, all patients who have been referred and fit the sample criteria are analysed by the protocol. The work-up according to the protocol as well as that by the cardiologist are noted and recorded. If in contrast with the opinion of the cardiologists a stress ECG is advised by the protocol, this will not be performed. If fluoroscopy is recommended by the protocol it is performed additionally if needed and informed consent has been given by the patient. The detection of coronary calcification by simple fluoroscopy is performed in the conventional way using Siemens radiographic equipment, consisting of a cesium iodide input phosphor 6.5 inch image intensifier connected to a TV monitor. The resolution of the TV monitor is 14 lp/cm.

To validate simple fluoroscopy as a predictor for coronary artery disease, the result will be compared with the outcome of arteriography. Because the diagnostic protocol is based on the findings of cine fluoroscopy it is necessary to verify whether the results of simple and cine fluoroscopy demonstrate conformity. Both tests are reported by one of two experienced radiologists and the referral for both tests was arbitrarily to one of these two radiologists. Even if not all non-invasive tests are performed as advised by the protocol, the post-test probability of disease of the patient can still be estimated by using all other available relevant information; using the same probability threshold of 15% a verdict on whether or not to refer for arteriography will also be obtained for these patients. The patient population used to evaluate the protocol was sampled during the period 1 November 1987 - 31 December 1988. Figure 12.1 demonstrates schematically the selection process of this population.

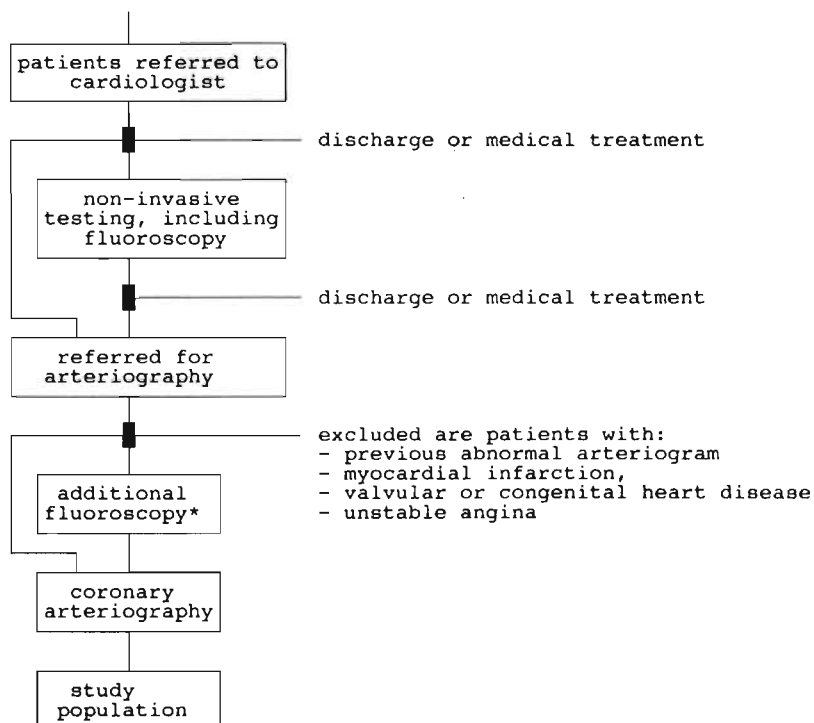


Figure 12.1 Selection process for the study population to evaluate the protocol

* As discussed in the text, dependent on the category of patients, it may occur that the protocol recommends fluoroscopy and that this examination has not yet been requested by the referring cardiologist

The data regarding to age, sex, symptoms, the results of the non-invasive tests performed and of the arteriogram will be collected and recorded in the same way as has already been described in Chapter 2. As this prospective study has two objectives, namely the evaluation of the protocol and the validation of simple fluoroscopy with cine fluoroscopy, the results with respect to these two issues will be addressed successively in the following.

12.2 Composition of the study population for evaluating the protocol: arteriographic findings, their distribution over gender, age and symptoms

During the period of November 1987 to December 1988, 582 patients were referred by the cardiologists for arteriography. After application of the exclusion criteria the study population consisted of 258 patients. The composition of this study population (in the following called the current population) is shown in Table 12.1, and is presented according to sex, age, and symptomatology and arteriographic findings combined.

Table 12.1 Composition of the current study population, consisting of 258 patients with (CAD+) or without (CAD-) significant coronary artery disease, categorised by sex, age, and symptoms and arteriographic findings combined

Sex	Age	Symptom*	CAD		
			CAD+	CAD-	Total
Male	< 50 yr	typical angina	21	3	24
		atypical angina	6	9	15
		n.c.p./asymptomatic	1	8	9
		Total	28	20	48
	50-59 yr	typical angina	20	3	23
		atypical angina	3	7	10
		n.c.p./asymptomatic	1	4	5
		Total	24	14	38
	≥ 60 yr	typical angina	43	8	51
		atypical angina	3	1	4
		n.c.p./asymptomatic	1	3	4
		Total	47	12	59
Female	< 50 yr	typical angina	3	4	7
		atypical angina	3	6	9
		n.c.p./asymptomatic	0	3	3
		Total	6	13	19
	50-59 yr	typical angina	13	7	20
		atypical angina	1	10	11
		n.c.p./asymptomatic	0	4	4
		Total	14	21	35
	≥ 60 yr	typical angina	25	12	37
		atypical angina	6	11	17
		n.c.p./asymptomatic	0	5	5
		Total	31	28	59

* Symptom as defined in Chapter 2

The protocol is based on data of 778 patients (base-line population) who consulted the cardiologists from 1983 to 1987. An essential presumption for the application of the protocol in this prospective study is that the association between the diagnostic variables and the disease status does not alter too much. Therefore, the distribution of the coronary arteriographic findings into 5 categories for both study populations is demonstrated in Table 12.2. The percentages in the categories are almost the same except in category 5 where the percentage in the current population is somewhat higher. This implies that the prevalence of patients with significant disease is higher in this population.

Table 12.2 Subdivision of arteriographic findings in 5 categories for the current and the base-line population

	Current population	Base-line population
Category 1*	26%	23%
Category 2*	3%	6%
Category 3*	13%	16%
Category 4*	7%	8%
Category 5*	52%	46%
Total	100% (N=258)	100% (N=778)

* Category 1: no coronary artery disease at all; Category 2: no coronary artery disease but only a muscle bridge; Category 3: only minimal coronary artery disease, no significant stenosis present; Category 4: coronary artery disease, with at least one lesion of 75% loss in cross sectional area (moderate stenosis) and the artery supplying 1.5 myocardial units, but with no lesions of 90%; Category 5: coronary artery disease, with at least one lesion of $\geq 90\%$ loss in cross-sectional area (severe stenosis), and the artery supplying 1.5 myocardial units

In Table 12.3 the distribution of the clinical variables is shown for both populations. Compared with the base-line population the current population consists of more females and there are relatively more patients older than 59 years at the expense of the age group of 50 to 60 years.

Table 12.3 Distribution of gender, age and symptomatology for the current and the base-line population

	Current population (N=258)	Base-line population (N=778)
Gender		
male	56%	68%
female	44%	32%
Age		
< 50 years	26%	29%
50-59 years	28%	44%
> 59 years	46%	28%
Symptomatology		
typical angina	63%	60%
atypical angina	26%	23%
n.c.p./asymptomatic	12%	17%

However, the most important item for the comparability of both study populations is the association between these clinical variables and the disease status. As discussed in Chapter 4 a proper way to express these associations are the ORs adjusted for collinearity, calculated with help of a logistic regression model. There are no important differences of the ORs for gender and symptomatology. In the current population the ORs for age are nearer to one, indicating that the association between age and disease status has become weaker (Table 12.4).

Table 12.4 Adjusted ORs for gender, age and symptomatology for the current and the base-line population with its 95%-Confidence Intervals in parentheses

	Current population (N=258)	Base-line population (N=778)
Gender		
male	1	1
female	0.32 (0.17-0.60)	0.24 (0.16- 0.36)
Age		
< 50 years	1	1
50-59 years	0.94 (0.42-2.13)	1.91 (1.23- 2.95)
> 59 years	1.45 (0.66-3.17)	3.29 (1.98- 5.43)
Symptomatology		
n.c.p./asymptomatic	1	1
atypical angina	5.75 (1.52-21.8)	3.94 (2.0 - 7.75)
typical angina	33.9 (9.33-123)	28.2 (15.0 -52.9)

12.3 Results of the application of the protocol

The application of the protocol to this study population of 258 patients would have reduced the number of referrals by 41. This comes to 34 investigations on an annual basis. In 12 of the 258 patients not all non-invasive testing was performed as advised by the protocol. This was the case 4 times for the stress ECG test, 6 times for fluoroscopy and in two patients both tests were lacking. The reason for missing out fluoroscopy was that the planning of the arteriogram was at too short notice. Thus, the protocol could be applied completely in 246 (258-12) patients.

For 213 of these 246 patients a post-test probability of 15% or more was estimated and consequently arteriography was advised by the protocol only for these patients (Table 12.5). From these 213 patients who were referred by the protocol no significant coronary artery disease was present in 65 patients. Thus, the false positive rate when applying the protocol, would have been 31% (65/213). As can be seen from Table 12.5 this percentage is 39% (97/246) for the cardiologists.

The post-test probability was estimated as less than 15% in 33 patients (Table 12.5) for whom the protocol recommends refraining from arteriography. However, significant coronary artery disease was present in one patient. According to the classification as used in Table 12.2 this patient was categorised as CAD 5 and underwent an elective PTCA procedure. The protocol advised correctly not to perform arteriography in 32 patients, 28 of whom had no disease at all, while 4

patients had only minor coronary artery disease (CAD 3). All 32 patients were discharged from the care of the cardiologist.

Table 12.5 Results of the protocol for patients with (CAD+) or without (CAD-) significant coronary artery disease for whom the protocol could be applied completely

	CAD+	CAD-	
Estimated post-test probability $\geq 15\%$	148	65	213
Estimated post-test probability $< 15\%$	1	32	33
	149	97	246

As stated, the protocol could not be applied completely in 12 patients, due to the lack of a stress ECG test, fluoroscopy or both, but although not all the necessary diagnostic information was available a post-test probability could still be estimated. For 4 patients the post-test probability was $\geq 15\%$, implying an indication for arteriography, and three of them proved to have significant coronary artery disease. For 8 patients the post-test probability was $< 15\%$ and consequently arteriography was not indicated. All of the latter eight patients proved to be free from significant disease. Adding these 12 patients to the 246 patients in whom the protocol could be applied completely, Table 12.6 can be constructed. From this table it can be seen that the false positive rate applying the protocol is still 31% but for the cardiologists this rate is now 42%.

Table 12.6 Results of the protocol applied for all patients with (CAD+) or without (CAD-) significant coronary artery disease

	CAD+	CAD-	
Estimated post-test probability $\geq 15\%$	149	68	217
Estimated post-test probability $< 15\%$	1	40	41
	150	108	258

A secondary effect of the application of the protocol is the change in indication for non-invasive testing. Table 12.7 gives an overview of the differences with the management of the cardiologists. For the stress ECG and thallium scintigraphy the protocol reduces the number of tests whereas it increases the number of the fluoroscopy tests indicated. In all of the 159 patients in whom the protocol advised against stress ECG testing, the cardiologists indicated that the stress ECG was performed for diagnostic reasons. Thus, the application of the protocol would have led to a saving of 159 stress ECG tests. As discussed in Chapter 11 there was no role allocated for thallium scintigraphy in the protocol. Therefore, the 42 thallium scintigrams that were performed during the period of November 1988 - December 1989, were assumed to be superfluous.

Using the same scale of charges of the Dutch Ziekenfondsen (semi-public, non-profit insurers which provide compulsory coverage for those earning less than Dfl. 50,000 a year) as explained in Appendix 1 of Chapter 11 the differences in indicated numbers of non-invasive tests can be expressed in Dutch guilders (Dfl.) as shown in the bottom row of Table 12.7. In total, the savings in non-invasive testing would amount to about Dfl. 50,774.

Table 12.7 Comparison of the numbers of non-invasive tests for 258 patients with symptoms suggestive of coronary artery disease as indicated by the protocol; with the numbers as indicated by the cardiologists

	Stress ECG	Fluoroscopy	Thallium scintigraphy
As indicated by the cardiologists	202	— *	42
As indicated by the protocol	49	82	0
Complying with the protocol as far as possible	43	74	
Difference in numbers	+ 159	— 74	+ 42
Difference in guilders (Dfl.)	+15,900	—925	+35,799

* The costs related to referrals for simple fluoroscopies by the cardiologists are not included, because at that moment these fluoroscopies were only requested to get an impression of the diagnostic value and not for the medical work-up of the patient

Apart from the risk of death or morbidity due to arteriography and apart from the inconveniences, this investigation generates costs. As explained in Appendix 1 of Chapter 11, the costs of an arteriogram have been set at Dfl. 2800. As application of the protocol would reduce the number of arteriographic investigations by 41, this amounts to a reduction of costs of Dfl. 114,800. Together with the savings on non-invasive testing the total saving would be Dfl. 165,574. However, on the debit side of the application of the protocol is the faulty withholding of an arteriogram from one patient. This was a 36 year old patient with atypical angina whose arteriogram demonstrated a 99% lesion in a large intermediate branch. The myocardial score was 3.0. This patient underwent a PTCA procedure.

It was expected that the protocol would incorrectly withhold arteriography from a number of patients. The problem is whether the number of false negative results was due to serious imperfections of the protocol or that they may be expected as a result of the chance mechanism inherent to any diagnostic protocol.

In our case there are 41 patients with an estimated probability of less than 15% and consequently the protocol advises that arteriography should not be carried out. The question that has to be answered, is what is the maximal number of false negative results that can be reasonably explained by chance, assuming that the probability estimates of the log-linear model are correct. The average post-test probability values of these 41 patients is 6.7% and therefore the point estimate of the number of false negatives for these 41 patients is 2.7 (0.067×41). Comparing this number of 2.7 with the number of one false negative patient found in reality

indicates that the protocol is at least cautious enough in withholding coronary arteriography.

However, to analyse this problem more completely it is necessary to determine a probability distribution for the different numbers of false negative results under the condition that the protocol is correct. This distribution is constructed by means of a Monte Carlo simulation technique.*

The results of this method are shown graphically in Figure 12.2. The null hypothesis that has to be tested is that the probability analysis on which the protocol is based is correct and that the observed number of false negative results can be explained by chance. This is done by calculating the probability of finding the observed number (or higher numbers) of false negative results assuming the correctness of the protocol (one-sided P-value). In our case the observed number is one. The probability of finding one or more false negatives is the sum of all columns less the first one in Figure 12.2 and equals 94.4%. In view of this value it may be argued whether the number of only one observed false negative result is too low with respect to the expected numbers and that the protocol is too cautious in its referral for arteriography.

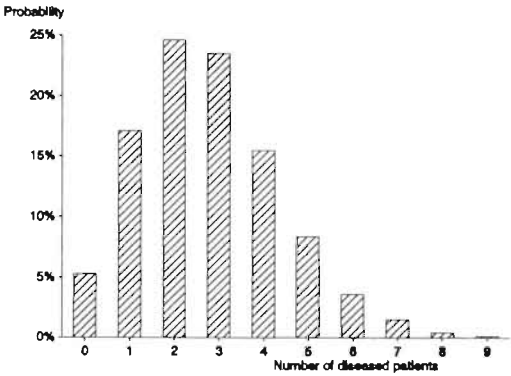


Figure 12.2 Distribution of the expected number of patients with coronary artery disease with a post-test probability < 15%, according to the log-linear model, as expected by the Monte Carlo simulation method after 100,000 simulations

The above mentioned procedure has also been used to test the practical implications of applying the protocol, particularly the number of false negative results. This approach is not appropriate to test the accuracy of the probability estimates of the log-linear model on which the protocol is based.

* The chance mechanism that generates the false negative results can be seen as a throw with 41 different dice. The dice are not the typical ones with 6 sides. Each of the dice has a different number of sides and therefore a different probability of showing the "black" side that corresponds with a patient's probability of having disease. For example, a patient with a disease probability of 12.5% could be represented by a dice with 8 sides of which one is black. With help of a computer program a great number of throws with these 41 dice can be simulated and for each throw the total number of black sides are counted and tallied. The distribution used in our analysis is constructed after 100,000 of those throws.

The model provides post-test probability estimates of having coronary artery disease for all categories of patients and these estimates can be compared with the observed values. However, for the majority of these categories the number of patients is low or even equal to zero, impairing a correct comparison. Therefore, the patients have to be amalgamated into categories with greater numbers. In our case the patients have been re-categorized according to their post-test probability estimates. The cut-off points for these categories have, rather arbitrarily, been set at 25%, 50% and 75%. With help of the chi-square test it is now possible to obtain an impression of the fit of the model (Table 12.8). It was ensured that each category contained at least 5 patients, a requirement for using this test.

As can be seen, the model, for instance, estimates for 101 patients a post-test probability that exceeds 75%. Further, it predicts that 91.1 of those 101 patients will have coronary artery disease. In reality, 87 patients proved to have disease. The other categories demonstrate differences as well. With help of the chi-square test it is possible to test if these differences can be explained by the chance mechanism. The p value of 41% (Table 12.8) implies that the mathematical model describes the data properly and that the difference in results between the model and the observed results can be ascribed to chance.

Table 12.8 The chi-square test, (χ^2)-goodness-of-fit test, for the log-linear model applied to the prospective population

Post-test probability range	Total number of patients	Actual number of patients with CAD	Number of patients with CAD as expected by the model
0- 24%	61	5	6.3
25- 49%	36	16	13.
50- 74%	60	42	38.7
75-100%	101	87	91.1
Total	258	150	149.1

CAD = significant coronary artery disease; $\chi^2_4 = 4.0$; $p = 0.41$

In conclusion, analysis of the results with help of the Monte Carlo simulation technique demonstrates that the observed number of false negative results is not alarming and the chi-square test demonstrates a proper fit of the log-linear model on which the diagnostic protocol is based.

12.4 Some remarks with respect to the interpretation of the results of the protocol

Objections which may be raised against the use of the protocol may include missing the diagnosis of left main disease, since this type of coronary artery disease is a life-threatening condition and bypass surgery improves life expectancy in these cases. In the current study population 12 patients (5%) proved to have left main disease. All these patients had typical angina and belonged to a category of patients for which the protocol advised arteriography without further non-invasive testing. (There were no patients with unstable angina

and left main disease because patients with unstable angina were not included in this study.) Patients with left main disease almost always have typical or unstable angina and except for females younger than 50 years the protocol advises arteriography without further testing. Therefore, the chance of missing left main disease is small.

The protocol advises against an arteriogram in patients with only minor disease. Thus, a second argument opposing the use of a protocol could be that the cardiologist wishes to detect these patients in order to treat them preventively. However, the 5 patients of the current population falling in this category were discharged from cardiological care. The 127 patients of the baseline population who fell also in this category were handled identically. Detecting patients with minor disease is apparently not too crucial.

A last issue with respect to the evaluation of the results of the protocol is the following.

There are patients for whom an arteriogram will be performed, although the referring physician is convinced that no abnormality exists. The rationale of such referral is that documenting the absence of significant coronary artery disease will reassure both the patient and the physician. As a result the patient can return to an active lifestyle without fear of cardiac disability. Thus, these patients undergo cardiac catheterisation despite a low probability of coronary artery disease based on clinical history and non-invasive testing.

As judged by our protocol 41 patients had no indication for coronary arteriography: 40 of them proved to have no significant disease, 35 patients had no coronary artery disease at all and 5 had only minor disease. It could be argued that proving the absence of disease by a coronary arteriogram would reduce medical consumption and/or would provide savings due to work resumption. There is, as judged from the literature, a difference of opinion about the value of a negative (normal) coronary arteriogram (Waxler, 1971^{W9}; Ockene 1980^{O2}; Faxon, 1982^{F5} and 1983^{F6}; Papanicolaou 1986^{P9} and Van Dorpe 1987^{D16}). In these studies a substantial number of patients continue to use medical facilities in spite of a negative arteriogram.

In selected cases there may be an indication for arteriography for reasons of reassurance. However, this referral indication is difficult to generalise and as a protocol always leaves room for individualised decisions such referrals do not conflict with the protocol.

12.5 Conclusion

The application of the protocol reduces substantially the number of referrals for coronary arteriography of patients without significant disease and non-invasive testing at costs that seem acceptable: one single patient with coronary artery disease would have been denied arteriography and the subsequently PTCA procedure. The arguments against the use of the protocol with respect to left main disease, minor disease and the reassurance referral can be refuted.

Thus, the questions, "Is application of a protocol justified?", or "Does it impair proper handling of the patients?" have been answered indisputably in the forgoing, moreover they should be replaced by the question "Is working without a protocol justified?".

13

Validation of simple fluoroscopy with cine fluoroscopy as method for detection of calcification in the coronary arteries

As discussed in Chapter 12 one of the reasons for performing a prospective study was to evaluate the performance of simple fluoroscopy as a non-diagnostic test to detect calcification in the coronary arteries.

Simple fluoroscopy was performed in 143 out of the 258 patients who took part in the evaluation of the protocol. The protocol advised this test in 82 patients whereas it was actually performed in 74; in 8 patients in whom simple fluoroscopy was also indicated by the protocol the test was not performed for reasons discussed in the previous chapter.

Additionally to these 74 tests, simple fluoroscopy was performed in 69 patients on account of a referral by the cardiologists.

Simple fluoroscopy as a predictor for coronary artery disease was validated with the result of arteriography.

Calcification was detected in one or more coronary arteries in 49 patients, 40 of whom had significant coronary artery disease (Table 13.1). 9 patients demonstrated calcifications but their vessels showed only minor stenoses or only wall irregularity. No calcification was demonstrated in the coronary arteries of 94 patients although significant disease was present in 25 and absent in 69.

As a parameter of the diagnostic value a stratified Odds Ratio was calculated with the help of logistic regression. The OR of simple fluoroscopy, stratified for gender, age and symptomatology, was 29.8.

Subdivision of the positive results for calcification according to the location of the calcification revealed that 30 patients showed calcification in only the left coronary artery, of whom 23 had significant disease. 19 patients had calcification in the right coronary artery of whom 17 had significant disease. The stratified OR for positive findings in at least the right coronary artery was 55.9 and the stratified OR for positive findings in only the left coronary artery was 21.5 (Table 13.1).

Calcification in only the right coronary artery occurred in two patients of whom both had significant coronary artery disease.

Furthermore, calcification was present at the site of a significant lesion according to the findings of arteriography in 30 of these 40 patients (75%). The remaining 10 patients showed only minor disease at the site of the calcification but had at least one significant lesion elsewhere in their coronary arterial tree.

The validation of the findings of calcifications by simple fluoroscopy was done with the findings of cine fluoroscopy.

Table 13.1 Presence of coronary artery calcification by simple fluoroscopy, specified for localisation, for patients with (CAD+) or without (CAD-) significant coronary artery disease (LCA = left coronary artery only; RCA = right coronary artery only or both coronary arteries) and the crude OR and the adjusted OR with its 95%-Confidence Interval for these findings

	CAD+	CAD-	OR crude	OR, adjusted with 95%-C.I.
Calc-	25 (38%)	69 (88%)	1	1
Calc+	40 (62%)	9 (12%)	12.3	29.8 (7.8-113)
- LCA	23 (35%)	7 (9%)	9.1	21.5 (2.0- 92)
- RCA	17 (26%)	2 (3%)	23.5	55.9 (7.8-402)
	65 (100%)	78 (100%)		

The simple fluoroscopy findings were confirmed by the cine (fluoroscopy) findings in 131 patients. There was a mismatch in 12 patients, 8.3% (Table 13.2a). Fluoroscopy over-diagnosed calcification in 1 patient but missed its presence completely in 6 patients and incompletely in 5 patients by only detecting the calcification in the left coronary artery tree but missing the calcification on the right side. For patient management the test is said to be positive if any calcification is detected. Thus, the findings of these 5 patients were also interpreted as a positive test result. The comparison of the test results of both forms of fluoroscopy, as used for patient management, is demonstrated in Table 13.2b. The Kappa as parameter for inter-test agreement is 0.90 with a 95%-Confidence Interval of 0.73-1.0.

Table 13.2a Comparison of the finding of calcification by simple fluoroscopy and cine fluoroscopy

		Cine fluoroscopy		
		+	-	
Fluoroscopy	+	48	1	49
	-	11	83	94
		59	84	143

Table 13.2b Comparison of the test results of simple fluoroscopy and cine fluoroscopy

		Cine fluoroscopy		
		+	-	
Fluoroscopy	+	53	1	54
	-	6	83	89
		59	84	143

For the analysis of these mismatches and to correlate them with the arteriographic findings, Table 13.3 was constructed. The single patient who demonstrated calcifications only according to the simple fluoroscopy proved to have completely normal coronary arteries. Calcification which proved to be present in the left coronary artery was missed by simple fluoroscopy in 6 patients, 4 of whom had significant disease. No special pattern connecting one radiologist to these mismatches could be demonstrated. The results of the comparison of simple fluoroscopy as non-invasive test with findings for cine fluoroscopy as detection mode of calcification show that the stratified OR for simple fluoroscopy test is 29.8 (95%-C.I. = 7.8-113) and for cine fluoroscopy is 28.0 (95%-C.I. = 7.9-100).

Table 13.3 Summary of mismatches in fluoroscopic and cine fluoroscopic detection of calcification in coronary arteries

	Simple Fluoroscopy	
	no calcification	calcification left tree
Cine fluoroscopy		
no calcification	-	1 (no CAD, cat. 3)
calcification left tree	2 (no CAD, cat. 3) 4 (CAD, cat. 5)	-
calcification both trees	-	5 (CAD, cat. 5)

CAD and category numbers are explained in Chapter 12

Conclusion

The disagreement between the results of simple fluoroscopy and those of the cine fluoroscopy is not alarming.

Calcification occurs more often in the left coronary artery tree. As the incomplete mismatches were caused only by missing calcifications in the right coronary artery tree, it seems that it is easier to detect calcification in the left than in the right tree.

As can be concluded by the value of the stratified OR, simple fluoroscopy is a worthwhile non-invasive test for patients suspected of coronary artery disease.

Its use, as judged by the protocol, is especially recommended for patients of both sexes with n.c.p./asymptomatic, females with atypical angina and females younger than 50 years with typical angina.

14.1 Aim of the thesis

As limitations on health care resources increase, it becomes essential to try to reduce the number of unnecessary investigations and to improve the cost effectiveness of the diagnostic process. Cine coronary arteriography is one of the most expensive invasive examinations in diagnostic radiology.

The aim of this study was to develop a method for evaluating and optimising the work-up in diagnostic radiology for a certain diagnostic problem, accounting for the specific situation of each patient. Probability analysis in diagnosing coronary artery disease serves as model.

In the department of diagnostic radiology at the De Wever Hospital, the radiologist is closely involved in cine coronary arteriography. Moreover, the reporting of the studies is standardised (Chapter 3). This facilitated the composition of a study population, enabling a study on the formulation of a protocol for the diagnostic work-up and referral for a cine coronary arteriogram.

Following a literature study, a study population ($n=778$) was compiled in order to compose such a protocol. This protocol was then prospectively applied to a new study population ($n=258$). Before performing the arteriogram, it was determined whether or not it would have been advised by the protocol.

The most important conclusions are summarised in the following sections.

14.2 Conclusions in regard to the method used for the probability analysis

The usual and easiest approach for a probability analysis is to use Bayes' theorem, while assuming independence between all diagnostic variables. It was not chosen as our method, as it may be questioned if this assumption is correct. Dependence consists of two phenomena, collinearity and interaction (Chapter 4).

Instead, a regression model was used for the probability analysis. It provides a sophisticated way to account for collinearity and can take interaction, if necessary, into consideration as well (Chapter 4).

A logistic regression model was used as an analytical tool for analysing the importance of dependence (Chapter 4). Using this regression model and accounting for dependence between the diagnostic variables, there appeared to be substantial differences in the posterior probabilities between the independent model and the method accounting for dependence (stress ECG, Chapter 7; fluoroscopy, Chapter 9; and combinations of the different non-invasive tests, Chapter 11). Dependence appeared to be of importance, as the differences in results with the two methods led also to differences in management of patients (Chapter 11).

For the definitive construction of the protocol a hierarchical log-linear model was chosen as regression method (Chapter 5). This type of regression model provides an estimate of the real number of patients with and without disease and not just an estimate of the ratio of the number of diseased to non-diseased patients, as is the case with a logistic regression model.

In conclusion, for the probability analysis of a diagnostic test or protocol, one needs to account for collinearity and to take interaction into consideration if necessary. A log-linear model as regression model is advised.

14.3 General conclusions regarding the protocol in the diagnosis of coronary artery disease

In this thesis a protocol is proposed for the diagnosis of coronary artery disease (Chapter 11). Sex, age and symptomatology are used as clinical diagnostic variables, resulting in a number of strata. Consequently an advice is given per stratum whether to fluoroscope the patient to detect coronary artery calcifications or to submit the patient to stress ECG testing. Depending on the results of these tests, an advice is given whether or not to perform a cine coronary arteriogram.

If a patient has $\geq 15\%$ chance of having coronary artery disease an arteriogram is advised. This norm is based on the budgetary possibilities of the De Wever Hospital and on the size of the patient population needing an arteriogram (Chapter 5).

Naturally, there are advantages and disadvantages of using a protocol. It is concluded that this protocol, taking cost-effectiveness into consideration, meets the present criteria for proper health care. Questions, "Is application of a protocol justified?", or "Does it impair proper handling of the patients?" should be replaced by the question "Is working without a protocol justified?" (Chapter 12).

A number of objections may be raised against the study. The protocol involves patients referred for cine coronary arteriography in the De Wever Hospital. Thus, the protocol only involves referred patients. This implies a certain patient selection, and these selected may differ from those consulting the cardiologist (Chapter 2).

The clinical history, the characterising of the complaint, forms a central issue in the protocol. However, history taking was not standardised (Chapter 6).

Furthermore, the stress ECG testing and the thallium scintigram testing form a problem. Both these tests were not performed in all patients of the study population (Chapter 5 and 8).

There are, of course other diagnostic variables such as blood cholesterol levels, family history, smoking habits, but these have not been included in this analysis. As judged by protocols in the cardiological literature these are not used, and do not seem to be the most important ones. The protocol constructed provides an advice based on the most important variables, but of course the clinician is free to change the management on the basis of other diagnostic information. A further differentiation of the symptom typical angina into more classes was not attempted for the same reasons, apart from the fact that this information was almost certainly not available for all patients (Chapter 2). Limitation of the number of variables studied improves our insight in the analysis and is also necessary to produce a workable protocol.

Nevertheless, the important result of this thesis is not the protocol for diagnosing coronary artery disease as such. It only served as a model to approach the final aim of this thesis, how a protocol for diagnostic radiology may be constructed.

If one does not agree with the budget and its consequences then a protocol provides an instrument to prove that the budget ought to be increased to provide the necessary medical care.

14.4 Some specific aspects of the protocol

The protocol does provide some important suggestions.

Every cardiologist will agree that proper history taking is crucial in diagnosing coronary artery disease (Chapter 6): careful history taking, especially of the symptomatology, usually provides a high degree of certainty whether or not coronary artery disease is present. This is reflected in the protocol; for example for males with typical angina an arteriogram is advised without any prior non-invasive testing. This implies a possible reduction in non-invasive testing.

Fluoroscopy for the detection of calcification in the coronary arteries plays an important role in the protocol. This non-invasive diagnostic test is not used in The Netherlands. The literature suggests that it may well be a useful test (Chapter 9).

Fluoroscopy is easy to perform, with additional advantages of being independent of medication and of the patient's exercise capacity, as well as having the benefit of relatively low costs. It can be performed in handicapped patients or in patients with a left bundle branch block.

All patients taking part in our study were fluoroscoped; the patients of the first study population were evaluated by the cine film; the patients of the second study population were fluoroscoped prior to the arteriogram if advised by the protocol. Analysis of the results for all these cases justified the central position of fluoroscopy in the protocol. Fluoroscopy proved to be very helpful especially in

patients who, according to their history, had a relatively low chance of having coronary artery disease (Chapter 9).

The stress ECG also plays an important role; however, its place is limited to specific categories of patients if fluoroscopy is performed, e.g. only males with nonspecific chest pain/asymptomatic and a negative fluoroscopy test (Figure 11.9, Chapter 11). In this protocol, intended specifically to confirm or exclude the presence of coronary artery disease, there appeared to be no reasons to include thallium scintigraphy.

14.5 Conclusion

In the coming years there will be mounting pressure on the physician from all who are involved in patient care (other health care professions, patients, hospital boards and insurance companies) to explain and justify the process of medical decision making. The obvious tool for this is the use of a protocol. In this study a method was demonstrated how to construct a diagnostic protocol based on empirical data of former patients, leading to suggestions for changes in patient management that seemed worthwhile. As patient management is dependent on many time-related factors, it is necessary to repeat such an exercise. By using continuously collected and valid data of diagnostic and clinical-outcome variables of patients, it is thus possible to periodically adapt the patient management.

Ontwikkelen en toetsen van een protocol ten behoeven van de diagnostiek van coronair vaatlijden

De financiële mogelijkheden in de gezondheidszorg worden steeds krappere en hebben geleid tot budgettering. Derhalve is het van steeds groter belang ondermeer het aantal onnodige diagnostische onderzoeken te reduceren en de kosteneffectiviteit van het diagnostische proces te verbeteren. Coronairarteriografie is nog steeds de enige methode om met "zekerheid" de diagnose en de uitgebreidheid van het vaatlijden vast te stellen, maar het is een van de kostbare invasieve diagnostische onderzoeken.

Bovendien blijkt de indicatiestelling tussen cardiologen en bij eenzelfde cardioloog in de loop van de tijd nog al te verschillen.

Ten behoeve van de indicatiestelling voor een arteriogram kan gebruik worden gemaakt van een inspannings ECG en thalliumscintigrafie naast anamnestiche gegevens. Verder wordt in de literatuur aangegeven dat ook de detectie van kalk in de coronair arteriën door middel van doorlichting (fluoroscopy) een zinvolle niet-invasieve test zou zijn.

De opzet van deze studie is het ontwikkelen van een protocol, dat aan dient te geven bij wie een arteriogram zou moeten worden verricht en welke niet-invasieve testen van te voren zouden moeten worden uitgevoerd.

Verder wordt de waarde van het doorlichtingsonderzoek naar kalkafzetting in de kransslagaderen nader geëvalueerd (hoofdstuk 1).

In hoofdstuk 2 worden doelstellingen, opzet en verzameling van gegevens van het onderzoek besproken.

In navolging van de literatuur werd voor de klinische variabelen leeftijd, sexe en symptomatologie gekozen en verder als diagnostische variabelen voor het inspannings ECG, de thalliumscintigrafie en de doorlichting op kalk.

De gegevens voor de samenstelling van het protocol kwamen voort uit een retrospectief onderzoek. Het bleek dat niet bij alle patiënten van de studiepopulatie (n=778) een inspannings ECG of een thalliumscintigram werd verricht. Dit is inherent aan een retrospectieve observationele studie. In hoofdstuk 2 wordt hierop ingegaan. Ook andere bezwaren verbonden aan een retrospectieve opzet worden in dit hoofdstuk besproken.

Indicatiestelling voor en risico's van coronairarteriografie zowel als definiëring van een laesie van haemodynamische betekenis komen in hoofdstuk 3 ter sprake. In

het De Wever Ziekenhuis te Heerlen is de radiodiagnost(e) nauw bij de uitvoering van cinecoronairarteriografie betrokken, de verslaggeving is gestandaardiseerd zoals beschreven door Brandt in 1977.

In hoofdstuk 4 worden de parameters voor het beschrijven van de discriminerende waarde van een diagnostische test, zoals sensitiviteit en specificiteit, beschreven. Er wordt uiteengezet waarom de voorkeur moet worden gegeven aan de "odds-ratio" als testparameter. In dit hoofdstuk wordt tevens ingegaan op de statistische afhankelijkheid tussen diverse diagnostische variabelen en de mogelijke invloed hiervan op de waarschijnlijkheidsanalyse. Statistische afhankelijkheid kan optreden via collineariteit en via interactie.

Een elegant model voor zo'n waarschijnlijkheidsanalyse, waarbij rekening wordt gehouden met deze statistische afhankelijkheid, is de logistische regressie.

Essentieel voor de constructie van het protocol is het ontwikkelen van een maatstaf ter bepaling of een patiënt al dan niet in aanmerking komt voor een arteriogram. De drempel waarboven wel een arteriogram wordt geadviseerd en waaronder geen arteriogram wordt geadviseerd kan op verschillende manieren worden vastgesteld, zoals door middel van de klassieke methode van een medische besliskundige analyse of op basis van consensus. In hoofdstuk 5 wordt besproken hoe ook zonder een uitgebreide medisch besliskundige analyse deze drempel ($P_{\text{threshold}}$) kan worden bepaald en een protocol kan worden opgesteld dat aangeeft:

- bij wie coronair arteriografie zou moeten worden verricht
- welke non-invasieve diagnostiek voorafgaand aan deze beslissing zinvol is, en
- wat de optimale volgorde is van deze non-invasieve diagnostiek.

In deze studie wordt bij de bepaling van de $P_{\text{threshold}}$ uitgegaan van het budget. Hierbij wordt rekening gehouden met de beschikbare middelen en het aantal patiënten dat een arteriogram zou moeten ondergaan.

Door middel van een hiërarchisch log-lineair model werd een cumulatieve lijst samengesteld van de patiënten op basis van de hoogte van de kans op coronair vaatlijden. De kans op ziekte werd bepaald op basis van de combinatie van de uitslagen van alle diagnostische variabelen. Gelet op het beschikbare budget leidt dit tot een $P_{\text{threshold}}$ van 15%. Volgens de literatuur is dit een redelijke waarde.

Zou in de toekomst blijken dat bij een op deze wijze vastgestelde $P_{\text{threshold}}$ de faciliteiten voor coronair arteriografie ondanks de scherpere selectie te kort schieten, en dat dus het budget te laag is, dan heeft men een waardevol instrument in handen om een aanvraag tot verhoging van het budget te onderbouwen (hoofdstuk 5).

De volgende hoofdstukken behandelen de relevante literatuur, zowel als onze eigen resultaten, over de waarde van de gebruikte klinische variabelen (hoofdstuk 6), het inspannings ECG (hoofdstuk 7), thalliumscintigrafie (hoofdstuk 8) en de (cine)doorlichting (hoofdstuk 9). Uit de resultaten blijkt dat in de waarschijnlijkheidsanalyse met statistische afhankelijkheid moet worden rekening gehouden. Tevens blijkt cine-doorlichting een zinvol onderzoek te zijn, vooral bij patiënten die anamnestic een geringe kans op coronair vaatlijden hebben.

In hoofdstuk 10 is de literatuur beschreven die protocollen voor de diagnostiek van coronair vaatlijden behandelt. Deze protocollen gaan evenwel allen uit van

onafhankelijkheid tussen de diagnostische variabelen. De populaties waarop deze protocollen berusten variëren sterk in samenstelling. De protocollen zijn ook vrij eenvoudig.

Bovendien zijn de gepubliceerde gegevens zo ontoegankelijk, dat een waarschijnlijkheidsanalyse op basis hiervan, wel rekening houdend met afhankelijkheid, niet mogelijk is; reden waarom is gekozen voor het opstellen van een protocol uit een eigen onderzoekspopulatie.

In hoofdstuk 11 wordt de constructie van diagnostische stroom diagrammen beschreven. Hier blijkt weer dat ook bij een combinatie van niet-invasieve onderzoeken met statistische afhankelijkheid rekening moet worden gehouden. In dit hoofdstuk wordt uitgewerkt welk niet-invasieve onderzoek of eventueel welke combinatie van meerdere niet-invasieve onderzoeken geïndiceerd is. Vervolgens wordt bepaald wat de beste volgorde zou moeten zijn van deze meerdere testen. Dit resulteert in een advies (protocol) over de te volgen strategie voor de verschillende categorieën van patiënten, ingedeeld naar leeftijd, geslacht en symptomatologie.

In het volgende hoofdstuk (12) wordt beschreven hoe dit protocol met behulp van de gegevens van 258 patiënten prospectief werd getoetst door de uitkomst van het management van de cardiologen te vergelijken met de uitkomst van het schaduwbeleid indien het protocol zou zijn toegepast. Het blijkt dat door middel van het protocol een betere selectie mogelijk is van patiënten met haemodynamisch significant coronair vaatlijden en dat er een vermindering van het aantal niet-invasieve onderzoeken optreedt.

De eindconclusie in dit hoofdstuk is dat het protocol voldoet aan de huidige criteria voor een goede gezondheidszorg en dat de vraag, "Is het toepassen van een protocol gerechtvaardigd?" of "Verhindert toepassing van een protocol het juiste beleid bij een patiënt?" dient te worden vervangen door de vraag "Is het werken zonder protocol wel gerechtvaardigd?".

Detectie van kalk zal in de praktijk gebeuren door middel van röntgen-doorlichting. In het retrospectieve onderzoek werd het uitgevoerd door middel van cine-doorlichting. In hoofdstuk 13 wordt röntgen-doorlichting vergeleken met cine-doorlichting. Uit de resultaten van de prospectieve studie blijkt de discriminerende waarde van röntgen-doorlichting ten opzichte van cine-doorlichting niet essentieel te verschillen. Detectie van kalk in coronairarteriën is een zinvol onderzoek.

Hoofdstuk 14 bevat een samenvatting en een aantal conclusies.

Elke cardioloog weet dat een nauwkeurig opgenomen anamnese belangrijk is bij het vaststellen van kransslagaderlijden: een anamnese van met name de symptomatologie geeft een hoge mate van zekerheid over het al of niet aanwezig zijn van coronair vaatziekte. Dit komt ook tot uitdrukking in het protocol.

Bij mannen met typische klachten van angina pectoris wordt bijvoorbeeld een arteriogram geadviseerd zonder enig voorafgaand niet-invasief onderzoek.

Doorlichten ter opsporing van kalkafzetting in de coronair arteriën speelt een belangrijke rol in het protocol. Dit niet-invasieve onderzoek wordt binnen Nederland uitsluitend in Heerlen uitgevoerd. De test is gemakkelijk uit te voeren, is onafhankelijk van medicamenten en inspanningstolerantie van de patiënt. De

kosten zijn laag. Analyse van de resultaten rechtvaardigt een centrale plaats van doorlichting op kalk in het protocol.

Het inspannings ECG speelt ook een belangrijke rol. Deze test heeft echter alleen zin bij speciale categorieën van patiënten na doorlichting op kalk, bijvoorbeeld bij mannen met specifieke klachten en een negatieve doorlichting.

In dit protocol, speciaal bedoeld om coronair vaatlijden uit te sluiten dan wel te bevestigen, wordt aan thalliumscintigrafie geen rol toebedeeld. Het blijkt dat de discriminerende waarde van de test, vergeleken met de beide andere niet-invasieve testen, laag is. Een duidelijk oordeel over de waarde van scintigrafie kan echter niet worden uitgesproken omdat blijkt dat het onderzoek vooral bij patiënten is uitgevoerd met een lage kans op vaatlijden, en dat er waarschijnlijk sprake is van een bepaalde selectie.

Er zijn natuurlijk meer diagnostische en/of prognostische variabelen. Maar op basis van eerder verrichte waarschijnlijkheidsanalyses in de cardiologische literatuur is het aantal niet verder uitgebreid. Het protocol vormt een richtlijn voor het gebruik van niet-invasieve onderzoeken en arteriografie gebaseerd op de belangrijkste klinische variabelen, maar laat onverlet de mogelijkheid om op basis van andere diagnostische informatie de beslissing aan te passen.

Het hoofdstuk wordt afgesloten met de conclusie dat in de komende jaren de druk op medici door alle partijen betrokken bij de gezondheidszorg (andere gezondheidswerkers, patiëntenverenigingen, directies van ziekenhuizen en verzekeringsorganen) zal toenemen om besluitvorming van medisch handelen te verduidelijken en te rechtvaardigen. Het geëigende middel is het gebruik van een protocol. In dit proefschrift wordt een methode aangegeven hoe een diagnostisch protocol zou kunnen worden opgesteld, gebaseerd op empirische gegevens van patiënten waarbij de diagnose uiteindelijk door middel van arteriografie werd gesteld.

Aangezien de behandeling van patiënten afhankelijk is van tijdgebonden factoren, is het noodzakelijk een protocol steeds aan te passen. Door het verzamelen en vervolgens gebruiken van steeds hernieuwde en betrouwbare gegevens van diagnostische zowel als uitkomstvariabelen is het mogelijk het beleid periodiek aan te passen.

16 | References

I am grateful to the authors of the following works and papers on which I have drawn freely.

List of references

A

1. Appleby JL. Why doctors must grapple with health economics. *Br Med J* 294: 326, 1987.
2. Acinapura AJ, Rose DM, Cunningham JN, et al. Coronary artery bypass in Septuagenarians. *Circulation* 78, II suppl.: 179, 1988.
3. Adams DF, Abrams HL. Complications of coronary arteriography: A follow-up report. *Cardiovasc. Radiol* 2: 89, 1979.
4. Ascoop CA, Simoons ML, Egmond WG, et al. Exercise test, history, and serumlipid levels in patients with chest pain and normal electrocardiogram at rest: Comparison to findings at coronary arteriography. *Am Heart J* 82: 609, 1971.
5. Ascoop CA, Robles de Medina EO. De betekenis van inspanningselektrocardiografie voor de diagnostiek van coronaria aandoeningen. *Ned T Geneesk* 123: 888, 1979.
6. Aldrich RF, Brensike JF, Battaglini JW, et al. Coronary calcifications in the detection of coronary artery disease and comparison with electrocardiographic exercise testing. *Circulation* 59: 1113, 1979.
7. Ascoop C, Klein B, Niemeyer M, et al. On the clinical value of Thallium-201 washout analysis in the detection of multiple jeopardized myocardial regions. *Int J Cardiol* 11: 305, 1986.
8. Abrams HL, Editor. *Coronary arteriography. A practical approach.* Little, Brown and Company, Boston, 1983.

B

1. Braunwald E. Unstable angina. *Circulation* 80: 410, 1988.
2. Brandt P W T, Partridge J B, Wattie W J. Coronary arteriography; Method of presenting of the arteriogram report and a scoring system. *Clin Radiol* 28: 361, 1977.

3. Bruschke AVG. The diagnostic significance of the coronary arteriogram. A study on its value in relation to other diagnostic methods. Thesis. Groningen, 1970.
4. Brook RH, Park RE, Winslow CM, et al. Diagnosis and treatment of coronary artery disease; comparison of doctors attitudes in the USA and the UK. *Lancet* I: 750, 1988.
5. Beentjes LB, Medische stralenbelasting in Nederland. NVS-Publication, nr.14, 1990.
6. Begg CB, McNeil BJ. Assessment of radiological tests: Control of bias and other design considerations. *Radiology* 167: 565, 1988.
7. Bayes T. An essay towards solving a problem in the doctrine of changes. *Philos Trans R Soc Lond* L111 53: 370, 1763. or *Biometrika* 45: 293, 1985.
8. Bartel AG, Chen JT, Peter RH, et al. The significance of coronary calcification detected by fluoroscopy. *Circulation* 49: 1247, 1974.
9. Borer JS, Kent KM, Bacharach SL, et al. Sensitivity, specificity and predictive value of radionuclide cineangiography during exercise in patients with coronary artery disease: comparison with exercise electrocardiography. *Circulation* 60: 572, 1979.
10. Bouter LM, van Dongen MCJM. Epidemiologisch onderzoek, opzet en interpretatie. Bohn Stafleu Van Loghem, Houten/Antwerpen, 1991.
11. Berger BC, Watson DD, Taylor GJ, et al. Quantitative Thallium-201 exercise scintigraphy for detection of coronary artery disease. *J Nucl Med* 22: 585, 1981.
12. Blankenhorn DH. Coronary calcification: a review. *Am J Med Sci* 242: 1, 1961.

C

1. Cameron EWJ, Walker WS. Coronary artery bypass surgery. *Brit Med J* 300: 1219, 1990.
2. Chaitman BR, Bourassa MG, Davis K, et al. Angiographic prevalence of high-risk coronary artery disease in patient subsets. (CASS) *Circulation* 64: 360, 1981.
3. Cohn Peter F, Editor. Diagnosis and therapy of coronary artery disease. Martinus Nijhoff Publishing, Boston, 1985.
4. Chaitman BR, Waters DD, Bourassa MG, et al. The importance of clinical subsets in interpreting maximal treadmill exercise test results: the role of multiple-lead ECG systems. *Circulation* 59: 560, 1979.
5. Carboni GP, Celli P, D' Ermo M, et al. Combined cardiac cinefluoroscopy, exercise testing and ambulatory ST-segment monitoring in the diagnosis of coronary artery disease; a report of 104 symptomatic patients. *Int J Cardiol* 9: 91, 1985.
6. Chaitman BR, Brevers G, Dupras G, et al. Diagnostic impact of Thallium scintigraphy and cardiac fluoroscopy when the exercise ECG is strongly positive. *Am Heart J* 108: 260, 1984.

D

1. Dunning AJ, Grinten van der TED, Dantzig van A, et al. Kiezen en delen. Advies in hoofdzaken van de commissie Keuzen in de zorg. 1991.
2. Davis K, Kennedy J W, Kemp H G, et al. Complications of coronary arteriography from the collaborative study of coronary artery surgery (CASS). *Circulation* 59: 1105, 1979.
3. Detrie KM, Wright E, Murphy M L, et al. Observer agreement in evaluating coronary angiograms. *Circulation* 52: 979, 1975.
4. DeRouen TA, Murray JA, Owen W. Variability in the analysis of coronary arteriograms. *Circulation* 55: 324, 1977.
5. Diamond GA, Forrester JS. Analysis of probability as an aid in the clinical diagnosis of coronary artery disease. *N Engl J Med* 300: 1350, 1979.
6. Detrano R, Salcedo EE, Hobbs RE, et al. Cardiac cinefluoroscopy as an inexpensive aid in the diagnosis of coronary artery disease. *Am J Cardiol* 57: 1041, 1986.
7. Detrano R, Yiannikas J, Salcedo EE, et al. Bayesian probability analysis: a prospective demonstration of its clinical utility in diagnosing coronary disease. *Circulation* 69: 541, 1984.
8. Detrano R, Leatherman J, Salcedo EE, et al. Bayesian analysis versus discriminant function analysis: their relative utility in the diagnosis of coronary artery disease. *Circulation* 73: 970, 1986.
9. Diamond GA, Forrester JS, Hirsch M, et al. Application of conditional probability analysis to the clinical diagnosis of coronary artery disease. *J Clin Invest* 65: 1210, 1980.
10. Diamond GA, Staniloff HM, Forrester JS, et al. Computer-assisted diagnosis in the noninvasive evaluation of patients with suspected coronary artery disease. *J Am Coll Cardiol* 1(2): 44, 1983.
11. Dunn RF, Kelly DT. Clinical indications for Thallium - 201 myocardial perfusion scanning. *Aust N.Z J Med* 12: 294, 1982.
12. Diamond GA. Correspondence. Exercise testing in asymptomatic young men. *N Engl J Med* 314: 579, 1986.
13. Diamond GA. Reverend Bayes' silent majority. An alternative factor affecting sensitivity and specificity of exercise electrocardiography. *Am J Cardiol* 57: 1175, 1986.
14. Diamond GA. Selection bias and the evaluation of diagnostic tests. (author's reply). *J Chron Dis* 39: 359, 1986.
15. Douglas PS. Editorial. Gender, cardiology, and optimal medical care. *Circulation* 74: 917, 1986.
16. Dorpe van A, Piessens J, Willems JL, et al. Unexplained chest pain with normal coronary arteriograms. A follow-up study. *Cardiology* 74: 436, 1987.

E

1. Enthoven AC. Shattuck lecture - Cutting cost without cutting the quality of care. *N Engl J Med* 298: 1229, 1978.
2. Ellestad MH, Savitz S, Bergdal D, et al. The false positive stress test. *Am J Cardiol* 40: 681, 1977.

3. Elamin MS, Boyle R, Kardash MM, et al. Accurate detection of coronary heart disease by new exercise test. *Br Heart J* 48: 311, 1982.
4. Eggen DA, Strong JP, McGill HC. Coronary calcification. Relationship to clinically significant coronary lesions and race, sex, and topographic distribution. *Circulation* 32: 948, 1965.
5. Eaker ED, Packard B, Wenger NK, et al. Coronary artery disease in women. *Am J Cardiol* 61: 641, 1988.

F

1. Fletcher RH, Fletcher SW, Wagner EH. Clinical epidemiology, the essentials. Second edition. Williams & Wilkins, Baltimore, 1988.
2. Fagan TJ. Nomogram for Bayes' theorem. *N Engl J Med* 293: 257, 1975.
3. Fortuin NJ, Weiss JL. Exercise stress testing. *Circulation* 56: 699, 1977.
4. Frink RJ, Achor RWP, Brown AL, et al. Significance of calcification of the coronary arteries. *Am J Cardiol* 26: 241, 1970.
5. Faxon DP, McGabe CH, Kriegel DE, et al. Therapeutic and economic value of a normal coronary angiogram. *Am J Cardiol* 73: 500, 1982.
6. Faxon DP, McGabe C. A normal coronary angiogram: is it therapeutic? *Int J Cardiol* 3: 360, 1983.

G

1. Gordon PR, Abrams C, Gash AK, et al. Percutaneous catheterization risk factors in left main coronary artery stenosis. *Am J Cardiol* 59: 1080, 1987.
2. Gould KL, Lipscomb K, Hamilton GW. Physiologic basis for assessing critical coronary stenosis. *Am J Cardiol* 33: 87, 1974.
3. Goodwin JF. The clinical approach - Cui Bono?. *Eur Heart J* 12: 751, 1991.
4. Goldschlager Nora. Use of the treadmill test in the diagnosis of coronary artery disease in patients with chest pain. *Ann Intern Med* 97: 383, 1982.
5. Goldschlager Nora, Selzer A, Cohn Keith. Treadmill stress tests as indicators of presence and severity of coronary artery disease. *Ann Intern Med* 85: 277, 1976.
6. Gould KL, Mullan N. Dynamic cardiac scanning. *J Nucl Med* 25: 1380, 1984.
7. Guiney TE, Pohost GM, McKusick KA, et al. Differentiation of false- from true-positive ECG responses to exercise stress by Thallium 201 perfusion imaging. *Chest* 80: 4, 1981.

H

1. Heiberg AN. The doctor in the twenty first century. *Br Med J* 295: 1602, 1987.
2. Harrison DG, White CW, Hiratzka LF, et al. The value of lesion cross-sectional area determined by quantitative coronary angiography in assessing the physiologic significance of proximal left anterior descending coronary arterial stenoses. *Circulation* 69: 1111, 1984.
3. Hung J, Chaitman BR, Lam J, et al. A logistic regression analysis of multiple noninvasive tests for the prediction of the presence and extent of coronary artery disease in men. *Am Heart J* 110: 460, 1985.

4. Hosmer D, Lemeshow S. Applied logistic regression. John Wiley & Sons, New York, 1989.
5. Hamby RI, Tabrah F, Wisoff BG, et al. Coronary artery calcification: clinical implications and angiographic correlates. *Am Heart J* 87: 565, 1974.
6. Hung J, Chaitman BR, Lam J, et al. Noninvasive diagnostic choices for the evaluation of coronary artery disease in women: a multivariate comparison of cardiac fluoroscopy, exercise electrocardiography and exercise Thallium myocardial perfusion scintigraphy. *J Am Coll Cardiol* 4: 8, 1984.
7. Hlatky MA, Mark DB, Harrell FE, et al. Rethinking sensitivity and specificity. *Am J Cardiol* 59: 1195, 1987.
8. Hlatky MA, Pryor DB, Harrell FE, et al. Factors affecting sensitivity and specificity of exercise electrocardiography. *Am J Med* 77: 64, 1984.
9. Hlatky MA. Evaluation of diagnostic tests. *J Chron Dis* 39: 357, 1986.

I

1. ICRP Publication 60. 1990 Recommendations of the International Commission on Radiological Protection. Pergamon Press, Oxford, 1990.
2. Isner JM, Kishe J, Kent KM, et al. Accuracy of angiographic determination of left main coronary arterial narrowing. *Circulation* 63: 1056, 1981.
3. Iskandrian AS, Hakki A. Thallium-201 myocardial scintigraphy. *Am Heart J* 109: 113, 1985.
4. Iskandrian AS, Segal BL. Value of exercise Thallium-201 imaging in patients with diagnostic and nondiagnostic exercise electrocardiograms. *Am J Cardiol* 48: 233, 1981.

J

1. Johnson MR. A normal coronary artery: What size is it? *Circulation* 86: 331, 1992.

K

1. Kirklin JW, Barratt-Boyes BG. Cardiac surgery. John Wiley & Sons, New York, page VII, 1986.
2. Kicken PJH, Janssen JHA, Michels HR, et al. Radiation exposure during cardiac catheterization procedures, IRPA 7, II, page 737, IRPA Congress Sydney, 1988.
3. Klocke FJ. Measurements of coronary flow reserve: defining pathophysiology versus making decisions about patient care. *Circulation* 76: 1183, 1987.
4. Killip CM. Twenty years of coronary bypass surgery. *N Engl J Med* 319: 366, 1988.
5. Knill-Jones RP. Diagnostic systems as an aid to clinical decision making. *Br Med J* 295: 1392, 1987.

L

1. Larson EB. Ignorance is not bliss: Knowledge, information, and the diagnostic technology problem. *Am J Roentgenol* 145: 1124, 1985.

2. Lieber A, Jorgens J. Cinefluorography of coronary artery calcification. Correlation with clinical arteriosclerotic heart disease and autopsy findings. *Am J Roentgenol* 86: 1063, 1961.
3. Loecker TH, Schwartz RS, Cotta CW, et al. Fluoroscopic coronary artery calcification and associated coronary disease in asymptomatic young men. *J Am Coll Cardiol* 19: 1167, 1992.

M

1. McNeil BJ, Varady PD, Burrows BA, et al. Cost-effectiveness calculations in the diagnosis and treatment of hypertensive renovascular disease. *N Engl J Med* 293: 216, 1975.
2. Maynard A. Logic in medicine: an economic perspective. *Br Med J* 295: 1537, 1987.
3. Martin CM, McConahay DR. Maximal treadmill exercise electrocardiography and cardiac hemodynamics. *Circulation* 46: 956, 1972.
4. McNeer JF, Margolis JR, Lee KL, et al. The role of the exercise test in the evaluation of patients for ischemic heart disease. *Circulation* 57: 64, 1978.
5. Maddahi J, Garcia EV, Berman DS, et al. Improved noninvasive assessment of coronary artery disease by quantitative analysis of regional stress myocardial distribution and washout of Thallium-201. *Circulation* 64: 924, 1981.
6. Margolis JR, Chen JTT, Kong Y, et al. the diagnostic and prognostic significance of coronary artery calcification. *Radiology* 137: 609, 1980.
7. McCarthy JH, Palmer FJ. Incidence and significance of coronary artery calcification. *Br Heart J* 36: 499, 1974.
8. Melin JA, Wijns W, Van Butsele RJ, et al. Alternative diagnostic strategies for coronary artery disease in women: demonstration of the usefulness and efficiency of probability analysis. *Circulation* 71: 535, 1985.
9. McNeil BJ, Keeler E, Adelstein SJ. Primer on certain elements of medical decision making. *N Engl J Med* 293: 211, 1975.
10. Macartney FJ. Diagnostic logic. *Br Med J* 295: 1325, 1987.

N

1. Naunheim KS, Kern MJ, McBride LR, et al. Coronary artery bypass surgery in patients aged 80 years and older. *Am J Cardiol* 59: 804, 1987.
2. National Radiological Protection Board, "A national Survey of doses to patients undergoing a selection of routine X-ray examinations in English hospitals", NRPB-R200, Chilton, Didcot, Oxon OX11 0RQ, 1988.
3. NRC Handelsblad. Kabinet verlaagt stralingsnormen radio-activiteit, 17 Maart 1991.
4. Neutze JM, White HD. What contribution has cardiac surgery made to the decline in mortality from coronary heart disease. *Br Med J* 294: 405, 1987.
5. Niemeyer MG. The clinical value of planar Thallium-201 myocardial perfusion scintigraphy. Thesis. Leyden, 1989.
6. Niemeyer MG, Ascoop CAPL, Cramer MJ, et al. De additionele waarde van kwantitatieve analyse van Thallium inspannings scintigrafie bij de detectie van meertaks coronarialijden. *Ned Tijdschr Cardiol* 6: 6, 1988.

O

1. Orr JS. New lessons from atomic bomb survivors. *Br Med J* 296: 1488, 1988.
2. Ockene IS, Shay MJ, Alpert JS, et al. Unexplained chest pain in patients with normal coronary arteriograms. *N Engl J Med* 303: 1249, 1980.

P

1. Pohost GM, Henzlova MJ. The value of Thallium-201 imaging. *N Engl J Med* 323: 190, 1990.
2. Patterson RE, Horowitz SF, Engl S, et al. Can exercise electrocardiography and Thallium - 201 myocardial imaging exclude the diagnosis of coronary artery disease ? *Am J Cardiol* 49: 1127, 1982.
3. Paulin S. Assessing the severity of coronary lesions with angiography. *N Engl J Med* 316: 1405, 1987.
4. Philbrick JT, Horwitz RI, Feinstein AR. Methodologic problems of exercise testing for coronary artery disease: groups, analysis and bias. *Am J Cardiol* 46: 807, 1980.
5. Pryor DB, Harrell FE, Scott Rankin J, et al. The changing survival benefits of coronary revascularization over time. *Circulation* 76 (suppl V): V-13, 1987.
6. Phibbs B, Fleming T, Ewy G, et al. Frequency of normal coronary arteriograms in three academic medical centres and one community hospital. *Am J Cardiol* 62: 472, 1988.
7. Proudfit WL, Shirey EK, Sones FM. Selective Cine coronary arteriography. *Circulation* 32: 901, 1966.
8. Patterson RE, Engl S, Horowitz SF, et al. Bayesian comparison of cost-effectiveness of different clinical approaches to diagnose coronary artery disease. *J Am Coll Cardiol* 4: 278, 1984.
9. Papanicolaou MN, Califf RM, Hlatky MA, et al. Prognostic implications of angiographic normal and insignificant narrowed coronary arteries. *Am J Cardiol* 58: 1181, 1986.

R

1. Ruijs JHJ. Radiodiagnostiek goed bekeken ? Thesis. Nymegen, 1985.
2. Reiber JHC, Booman F, Gerbrands JJ, et al. Methode voor objectieve en kwantitatieve analyse van coronair cine angiogrammen. *Hart Bulletin* 12: 19, 1981.
3. Reiber JHC, Serruys PW, Editors. State of the art in quantitative coronary arteriography. Martinus Nijhoff, Dordrecht, The Netherlands, 1986.
4. Rijneke RD, Ascoop CA, Talmon JL. Clinical significance of up-sloping ST segments in exercise electrocardiography. *Circulation* 61: 671, 1980.
5. Rifkin RD, Hood WB. Bayesian analysis of electrocardiographic exercise stress testing. *N Engl J Med* 297: 681, 1977.
6. Raffo JA, Luksic IY, Kappagoda CT, et al. Diagnostic value of routine exercise testing in hospital patients with angina pectoris, *Br Med J* 2: 295, 1979,
7. Rifkin RD, Parisi AF, Folland E. Coronary calcification in the diagnosis of coronary artery disease. *Am J Cardiol* 44: 141, 1979.

8. Ross R, Glomset JA. The pathogenesis of atherosclerosis. *N Engl J Med* 295: 369, 1976.
9. Rothman KJ. *Modern epidemiology*. Little, Brown and Company, Boston, 1986.

S

1. Sackett DL, Haynes RB, Tugwell P. *Clinical epidemiology. A basic science for clinical medicine*. Little, Brown and Company, Boston, 1985.
2. Satler LF, Green CE, Wallace BB, et al. Coronary artery disease in the elderly. *Am J Cardiol* 63: 245, 1989.
3. Stadius ML, Alderman EL. Coronary artery revascularization. *Circulation* 82: 2231, 1990.
4. Steingart RM. In *The Heart Arteries and Veins*, J W Hurst, Editor in Chief. McGraw-Hill, Information Services Company, New York, page 357, 1990.
5. Simoons ML. Computer assisted interpretation of exercise electrocardiograms. Thesis. Utrecht, 1976.
6. Soufer R, Zaret BL. Nuclear cardiology. In Cohn PF, Editor. *Diagnosis and therapy of coronary artery disease*. Martinus Nijhoff, Boston, page 191, 1985.
7. Snellen HA. *De rontgenologische diagnostiek der sclerose van hart en aorta*. Thesis. Leyden, 1939.

T

1. Timmis AD. *Probability analysis in the diagnosis of coronary artery disease*. Editorial. *Br Med J* 291: 1443, 1985.
2. Turner DA. An intuitive approach to Receiver Operating Characteristic Curve analysis. *J Nucl Med* 19: 213, 1987.
3. Tellingan van C, Ascoop CA, Rijnke RD. On the clinical value of conventional and new exercise electrocardiographic criteria: a comparative study. *Int J Cardiol* 5: 689, 1984.
4. Thwaites BC, Quyyumi AA, Raphael MJ, et al. Comparison of the ST/heart rate slope with the modified Bruce exercise test in the detection of coronary artery disease. *Am J Cardiol* 57: 554, 1986.
5. Turner DA, Battle WA, Deshmukh H, et al. The predictive value of myocardial perfusion scintigraphy after stress in patients without previous myocardial infarction. *J Nucl Med* 19: 249, 1978.
6. Tanenbaum SR, Kondon GT, Veselik KE, et al. Detection of calcific deposits in coronary arteries by ultrafast computed tomography and correlation with angiography. *Am J Cardiol* 63: 870, 1989.
7. Tavel ME, Enas NH, Woods JR. Sensitivity and specificity of tests: Can the "silent majority" speak? *Am J Cardiol* 60: 1167, 1987.

U

1. Uretsky BF, Rifkin RD, Sharma SC, et al. Value of fluoroscopy in the detection of coronary stenosis: influence of age, sex, and number of vessels calcified on diagnostic efficacy. *Am Heart J* 115: 323, 1988.

V

1. Vogel RA, Kirch DL, Lefree MT, et al. Thallium-201 myocardial perfusion scintigraphy results of standard and multi-pinhole tomographic techniques. *Am J Cardiol* 43: 787, 1979.

W

1. Willerson JT. Selection of patients for coronary arteriography. *Circulation* 72, suppl V: V-3, 1985.
2. White CW, Wright CB, Doty DB, et al. Does visual interpretation of the coronary arteriogram predict the physiologic importance of a coronary stenosis? *N Engl J Med* 310: 819, 1984.
3. Weiner DA, Ryan TJ, McCabe CH, et al. Correlations among history of angina, ST segment response and prevalence of coronary artery disease in the coronary artery surgery study (CASS). *N Engl J Med* 301: 230, 1979.
4. Wall van der EE. Nucleaire cardiologie. *Hart Bulletin* 19: 47, 1988.
5. Watson DD, Campbell NP, Read EK, et al. Spatial and temporal quantitation of plane Thallium myocardial images. *J Nucl Med* 22: 577, 1981.
6. Wackers FJTh, Fetterman RC, Mattera JA, et al. Kwantitatieve analyse van Thallium-201 inspanningsscintigrafie. *Hart Bulletin* 17: 4, 1986.
7. Woodring JH, West JW. Coronary artery calcification by CT in patients over forty years of age. *Australas Radiol* 33: 79, 1989.
8. Weintraub WS, Madeira SW, Bodenheimer MM, et al. Critical analysis of the application of Bayes' theorem to sequential testing in the noninvasive diagnosis of coronary artery disease. *Am J Cardiol* 54: 43, 1984.
9. Waxler EB, Kimbiris D, Dreifus LS. The fate of women with normal coronary arteriograms and chest pain resembling angina pectoris. *Am J Cardiol* 28: 25, 1971.

Z

1. Zir LM, Miller SW, Dinsmore RE, et al. Interobserver variability in coronary angiography. *Circulation* 53: 627, 1976.
2. Zijlstra F, Van Ommeren J, Reiber JHC, et al. Does the quantitative assessment of coronary artery disease dimensions predict the physiologic significance of a coronary stenosis? *Circulation* 75: 1154, 1987.

Acknowledgement

Het idee om met de gegevens van verrichte coronairarteriogrammen iets te doen en in een proefschrift te bewerken werd op 21 oktober 1985 voor het eerst in een gesprek met een van mijn collega's, Dr J.J.H. Lamers, verwoord. Dat was het begin. Het resultaat ligt voor U.

Prof. Dr F.Sturmans, promotor,
Beste Ferdinand,

Artsen in de gezondheidszorg werkend, specialisten in ziekenhuizen, ontbreekt het over het algemeen aan voldoende kennis tot het verantwoord verzamelen en bewerken van patiënten gegevens. Het zelfde geldt voor de specifieke kennis over waarschijnlijkheidsanalyses. Je hebt altijd de mening geuit dat ervaring van bijvoorbeeld een vakgroep epidemiologie ook ten goede moet komen aan klinisch onderzoek resulterend in zinvolle en praktisch bruikbare resultaten.

Zo was je instelling toen ik je de eerste maal bezocht in Maastricht, op 9 september 1986, wel weer lang geleden. Je had toen al een zeer positieve instelling. Ik ben altijd blij geweest met je woorden van steun. Ik ben je zeer dankbaar voor al je hulp.

Ook wil ik graag je echtgenote danken voor de gastvrijheid bij jullie thuis.

In het najaar van dat zelfde jaar ontmoette je mijn leermeester in de cardio-radiodiagnostiek, Dr Peter Brandt uit Auckland, Nieuw Zeeland. Jullie overlegden samen en zouden mij samen bijstaan. Peter overleed in september 1991 aan een slopende ziekte. Hij heeft nog stukken van het proefschrift gelezen en kunnen corrigeren.

Ik ben aan hem grote dank verschuldigd, hij bracht op mij de liefde over voor het hart radiologisch gezien. Alles wat ik weet op cardio-radiologisch gebied heb ik van hem geleerd. Hij was niet alleen een begaafd didact maar ook een voorbeeld als mens.

Vervolgens introduceerde je, Ir F.Kessels, arts, verbonden aan de afdeling epidemiologie van de Universiteit Limburg, die toen juist op je afdeling was gestart met zijn werkzaamheden.

Best Fons, in december, kwam je langs in Heerlen en zagen wij elkaar voor het eerst.

Het is niet in woorden uit te drukken wat jij hebt bijgedragen aan dit proefschrift. Door je specifieke kunde en inzet is het geworden wat het is. Als je mij eerlijk vraagt wat een van je grootste gaven is dan is het wel je belangeloze inzet.

Hoe vaak hebben we wel niet samen gezeten in mijn kamer in Heerlen, ontelbare malen. Fons ik ben je grote dank verschuldigd. Ik hoop dat we ook in de toekomst samen bezig blijven. Tiny ook veel dank voor al de tijd gegeven.

Prof. Dr J.M.A.van Engelshoven, promotor,
Beste Jos,

De radiodiagnost zal steeds meer mee moeten denken over de juiste en optimale diagnostische benadering (en eventuele interventie) van een klinisch probleem. Het is fijn dat jij als radiodiagnost hebt gedacht over de essentie van dit proefschrift en dat je dat hebt gesteund met woord en daad. Zo'n positieve instelling is een groot goed. Heel veel dank. Ik hoop nog lang met jou en je afdeling hier in Limburg samen te werken in het belang van beiden afdelingen, ook voor alle assistenten radiodiagnostiek die we in beide ziekenhuizen mogen opleiden.

Dr Warren Smith, cardiologist at Greenlane Hospital, Auckland, en Dr A.C. Ascoop, cardiologist at St Antonius Ziekenhuis in Nieuwegein.

Dear Warren and beste Carl, you have both corrected parts of the thesis, the sections on the stress ECG and the thallium scintigraphy. I am very grateful for your expertise and help.

Ir P.J.H.Kicken, verbonden als klinisch physicus aan de afdeling radiodiagnostiek van het Academisch Ziekenhuis in Maastricht, ook U wil ik danken voor uw expertise aangaande het deel handelend over het Rontgen stralen risico in de cine coronair arteriografie. Het blijft toch nog altijd een ingewikkelde materie vooral indien men onderzoeken onderling wil vergelijken.

F.Bongaerts, assistent in opleiding tot radiodiagnost, werkend in het De Wever Ziekenhuis,

Beste Fons,

Jij was een van de eerste "buitenstaanders" die al lezend een indruk kreeg wat het proefschrift inhield. Je houding was zeer enthousiasmerend. Ook zeker dankzij jou heeft een aantal hoofdstukken de vorm gekregen waar ze uiteindelijk in gegoten zijn. Ook wij zaten vaak samen te werken. Een zeer plezierige eigenschap van jou kwam zo boven water: af is af; dus morgen gaan we snel weer verder. Heel veel dank voor al je hulp. Dat ook wij nog vaak samen mogen werken.

Dr Leyva, now at the Department of Oncology of the Childrens Mercy Hospital in Kansas City, Missouri, the thesis was already quite far advanced when you helped me with corrections of the style and in the English. Although we have never met even when you lived and worked in the Netherlands, our cooperation was always excellent. You gave me a rapid and efficient service, for which I thank you very much.

Dr Leigh Warner, radiologist, in Auckland,

Dear Leigh, when I visited you in December 1991, you said you would be glad to help me by reading the manuscript for the thesis and would help in correcting faults in language and readability. This was a great support. We studied together in 1976 for the College exams, and radiology brought us together again on the thesis. Many thanks for your time and energy, and also for Ann's support.

De maatschap radiodiagnostiek van het De Wever Ziekenhuis, Jullie hebben altijd achter het proefschrift gestaan, in woord en daad. Dat was heel belangrijk en fijn om te weten.

Ineke, mijn partner in de cardio-radiodiagnostische werkzaamheden, jou wil ik speciaal bedanken. Ten eerste voor je interesse en accuratesse in het beschrijven van cinecoronair arteriogrammen en ten tweede voor het feit dat je het altijd mogelijk maakte dat ik nog even thuis achter de computer bezig kon zijn.

De maatschap cardiologie,

Kragten, Lunde, Tjan, Van den Berg, Vermaat en Vincent, jullie waren zondermeer de helpers van het eerste uur, door me te behulpzaam te zijn bij het verzamelen van de patiënten gegevens. Wel bedankt.

De dames Clignet, Wetzels en Werts, van de medische bibliotheek van het De Wever Ziekenhuis,

jullie stonden altijd klaar. Jullie zorgden dat ik steeds weer de beschikking had over de relevante literatuur. Bedankt voor jullie altijd weer vrolijke inzet.

De dames werkend op de polikliniek cardiologie,

indien jullie niet de statussen van ver over de 1000 patiënten hadden aangereikt zou ik niet eens met het onderzoek hebben kunnen starten. Wel bedankt.

De heer T.Aarts, verbonden aan de afdeling epidemiologie van de Universiteit Limburg en Dr B. Davies als bacterioloog verbonden aan het De Wever Ziekenhuis hebben mij ook geholpen. Beiden maakten de tekst klaar voor de drukker. Jullie beiden bedankt.

Bij het bewerken van een proefschrift is een computer niet meer weg te denken. Zeer gaarne wil ik de heer C. Gielissen bedanken voor zijn inzet. Tot 's avonds laat kon ik altijd terecht als er een probleem was met de computer. U, en uw echtgenote, zeer bedankt voor al uw tijd.

Tot slot de beoordelingscommissie van het proefschrift bestaande uit Prof. Dr P.J.Brombacher, Dr F.W.H.M.Bar, Prof. Dr J.M.J.P.van der Linden, Prof. Dr O.C.K.M.Penn en Prof. Dr J.H.J.Ruys. U allen wil ik graag danken voor Uw tijd om dit proefschrift aan een kritisch onderzoek te onderwerpen.

Paul, jij de voorzitter, wist al lang dat ik met dit proefschrift bezig was. Jij hebt altijd warme belangstelling getoond. Je had altijd een luisterend oor. Het is een vreugde dat jij er op deze wijze bij betrokken bent.

In het laatste stadium hebben drie collegae radiodiagnosten, Dr G. Landman, Dr K.Schuur en Dr H.de Valois, het proefschrift ook nog bestudeerd. Beste Guido, Klaas en Hans, ook jullie zeer bedankt voor de bereidheid, hulp en tijd.

Tot mijn verdriet hebben een aantal belangrijke personen in mijn leven de afronding van dit proefschrift niet meer meegemaakt.

Ondermeer denk ik dan aan mijn vader die eerder overleed. Hij heeft altijd zijn grote interesse getoond in het proefschrift.

Mijn Moeder is mij tot zeer grote steun geweest en heeft altijd achter het proefschrift gestaan. Zij weet wat het "gekost" heeft. Helaas is ze te oud en te

verzwakt om hier op de promotie zelf aanwezig te zijn. Ze zal hier zeker in gedachten aanwezig zijn. Lieve Moeder, ik ben je grote dank verschuldigd.

Ik ben heel blij dat mijn beide kinderen, Amelia en Jurjen de afronding van het proefschrift en het moment van de promotie wel kunnen meemaken.

Amelia en Jurjen,

jullie weten wat het impliceert een proefschrift schrijven. Ik draag het gaarne ook aan jullie beide op.

I am grateful to my friends and colleagues for sharing their time and knowledge with me thereby helping unwittingly to formulate my opinions.

The help of all these persons undoubtedly prevented many errors. Any that remain are solely my responsibility.

6416 EL Heerlen, March 1993

Molenberglaan 69

Curriculum vitae

27 oktober 1939	geboren te 's-Gravenhage
1958	eindexamen HBS-B, Rijnlands Lyceum te Wassenaar medische studie Rijksuniversiteit Groningen Stichting Klinisch Hoger Onderwijs Rotterdam
april 1966	arts examen
1966 - 1967	militaire dienst
1968 - 1970	werkend in het kader van ontwikkelings hulp in Kenya
1973 - 1976	opleiding tot radiodiagnost in Auckland, Nieuw Zeeland, lidmaatschap Royal Australasian College of Radiologists
1976 - 1977	fellowship cardiac radiology Dr P.W.T. Brandt Greenlane Hospital, Auckland
1977	fellowship vascular radiology Prof. Dr E.K. Lang X-ray department Louisiana State University, New Orleans, Louisiana, USA
1978	afdeling radiodiagnostiek Dr D. Limburg, St Geertruiden Ziekenhuis, Deventer
mei 1978	ingeschreven in register SRC

Sinds 1978 verbonden aan het De Wever Ziekenhuis in Heerlen, werkend in de
maatschap radiodiagnostiek bestaande uit:

Dr J.J.H.Lamers, Dr G.Goedhard, Dr J.TH.K.G.N.Nievelstein, G.O.Veldhuyzen van
Zanten *, C.W.M.Versteeg *, Mevrouw J.L.Schreutelkamp *, H.A.Engelshove *,
Dr R.Goei *, I.H.de Ploeg *, en H.F.Odink * (* huidige maatschapsleden).

